

P1 1201352

REC'D 03 AUG 2004

WIPO

PCT

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office

July 30, 2004

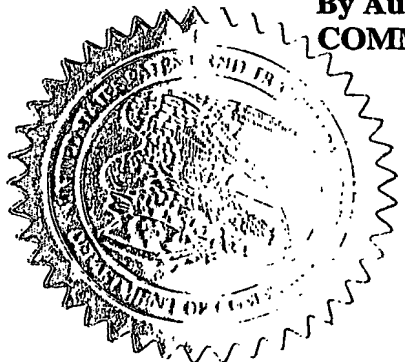
THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM
THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK
OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT
APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A
FILING DATE.

APPLICATION NUMBER: 60/566,212

FILING DATE: April 28, 2004

RELATED PCT APPLICATION NUMBER: PCT/US04/16614

By Authority of the
COMMISSIONER OF PATENTS AND TRADEMARKS



M. Tarver

M. TARVER
Certifying Officer

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

02570 U.S. PTO

2804

PTO/SB/16 (08-03)

Approved for use through 07/31/2008. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

Express Mail Label No. EU531571267US

INVENTOR(S)					
Given Name (first and middle [if any])		Family Name or Surname		Residence (City and either State or Foreign Country)	
Soumitra		Roy		Wayne, Pennsylvania	
Additional inventors are being named on the <u>1</u> separately numbered sheets attached hereto					
TITLE OF THE INVENTION (500 characters max)					
Methods of Generating Chimeric Adenoviruses and Uses For Such Chimeric Adenoviruses					
Direct all correspondence to: CORRESPONDENCE ADDRESS					
<input checked="" type="checkbox"/> Customer Number: 00270					
OR					
<input type="checkbox"/> Firm or Individual Name		Howson and Howson			
Address		Cathy A. Kodroff			
Address		Spring House Corporate Center, 321 Norristown Road, Suite 200			
City		Spring House		State	PA
Country		USA		Zip	19477
		Telephone	215-540-9200	Fax	215-540-5818
ENCLOSED APPLICATION PARTS (check all that apply)					
<input checked="" type="checkbox"/> Specification Number of Pages <u>65</u>		<input type="checkbox"/> CD(s), Number _____			
<input checked="" type="checkbox"/> Drawing(s) Number of Sheets <u>2</u>		<input checked="" type="checkbox"/> Other (specify) <u>159 pp. Sequence Listing</u>			
<input type="checkbox"/> Application Date Sheet. See 37 CFR 1.76					
METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT					
<input type="checkbox"/> Applicant claims small entity status. See 37 CFR 1.27.				FILING FEE Amount (\$)	
<input checked="" type="checkbox"/> A check or money order is enclosed to cover the filing fees.				<div style="border: 1px solid black; width: 100px; height: 100px; text-align: center; line-height: 100px;">160.00</div>	
<input checked="" type="checkbox"/> The Director is hereby authorized to charge filing fees or credit any overpayment to Deposit Account Number: <u>08-3040</u>					
<input type="checkbox"/> Payment by credit card. Form PTO-2038 is attached.					
The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.					
<input checked="" type="checkbox"/> No.					
<input type="checkbox"/> Yes, the name of the U.S. Government agency and the Government contract number are: _____					

15535 U.S. PTO
60/566212

042804

Respectfully submitted,

SIGNATURE Cathy A. Kodroff

TYPED or PRINTED NAME Cathy A. Kodroff

TELEPHONE 215-540-9200

[Page 1 of 2]

Date April 28, 2004

REGISTRATION NO. 33,980

(if appropriate)

Docket Number: UPN-P3067-P2USA

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

This collection of information is required by 37 CFR 1.51. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 8 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Mail Stop Provisional Application, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

PROVISIONAL APPLICATION COVER SHEET
Additional Page

PTO/SB/18 (08-03)
Approved for use through 07/31/2008. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

Docket Number UPN-P3067-P2USA

INVENTOR(S)/APPLICANT(S)		
Given Name (first and middle (if any))	Family or Surname	Residence (City and either State or Foreign Country)
James M.	Wilson	Gladwyne, Pennsylvania

[Page 2 of 2]

Number 1 of 1

WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.

02570 U.S. PTO
042604

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

Approved for use through 07/31/2008. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

FEE TRANSMITTAL for FY 2004

Effective 10/01/2003, Patent fees are subject to annual revision.

☐ Applicant claims small entity status. See 37 CFR 1.27

TOTAL AMOUNT OF PAYMENT (\$ 160.00)

Complete if Known

Application Number	
Filing Date	Apr 28, 2004
First Named Inventor	Roy et al
Examiner Name	
Art Unit	
Attorney Docket No.	UPN-P3067-P2USA

METHOD OF PAYMENT (check all that apply)

☒ Check ☐ Credit card ☐ Money Order ☐ Other ☐ None

☒ Deposit Account:

Deposit Account Number	08-3040
Deposit Account Name	Howson and Howson

The Director is authorized to: (check all that apply)

☐ Charge fee(s) indicated below ☒ Credit any overpayments
☒ Charge any additional fee(s) or any underpayment of fee(s)
☐ Charge fee(s) indicated below, except for the filing fee to the above-identified deposit account.

FEE CALCULATION

1. BASIC FILING FEE

Large Entity Fee Code (\$)	Small Entity Fee Code (\$)	Fee Description	Fee Paid
1001 770	2001 385	Utility filing fee	
1002 340	2002 170	Design filing fee	
1003 530	2003 265	Plant filing fee	
1004 770	2004 385	Reissue filing fee	
1005 180	2005 80	Provisional filing fee	160.00
SUBTOTAL (1)			(\$ 160.00)

2. EXTRA CLAIM FEES FOR UTILITY AND REISSUE

Total Claims		-20** =		X		=	
Independent Claims		-3** =		X		=	
Multiple Dependent						=	

Large Entity	Small Entity	
Fee Code (\$)	Fee Code (\$)	Fee Description
1202 18	2202 9	Claims in excess of 20
1201 86	2201 43	Independent claims in excess of 3
1203 290	2203 145	Multiple dependent claim, if not paid
1204 86	2204 43	** Reissue independent claims over original patent
1205 18	2205 9	** Reissue claims in excess of 20 and over original patent

**or number previously paid, if greater. For Reissues, see above

3. ADDITIONAL FEES

Large Entity Fee Code (\$)	Small Entity Fee Code (\$)	Fee Description	Fee Paid
1051 130	2051 65	Surcharge - late filing fee or oath	
1052 50	2052 25	Surcharge - late provisional filing fee or cover sheet	
1053 130	1053 130	Non-English specification	
1812 2,520	1812 2,520	For filing a request for ex parte reexamination	
1804 920*	1804 920*	Requesting publication of SIR prior to Examiner action	
1805 1,840*	1805 1,840*	Requesting publication of SIR after Examiner action	
1251 110	2251 55	Extension for reply within first month	
1252 420	2252 210	Extension for reply within second month	
1253 950	2253 475	Extension for reply within third month	
1254 1,480	2254 740	Extension for reply within fourth month	
1255 2,010	2255 1,005	Extension for reply within fifth month	
1401 330	2401 165	Notice of Appeal	
1402 330	2402 165	Filing a brief in support of an appeal	
1403 290	2403 145	Request for oral hearing	
1451 1,510	1451 1,510	Petition to institute a public use proceeding	
1452 110	2452 55	Petition to revive - unavoidable	
1453 1,330	2453 665	Petition to revive - unintentional	
1501 1,330	2501 665	Utility issue fee (or reissue)	
1502 480	2502 240	Design issue fee	
1503 640	2503 320	Plant issue fee	
1460 130	1460 130	Petitions to the Commissioner	
1807 50	1807 50	Processing fee under 37 CFR 1.17(q)	
1808 180	1808 180	Submission of Information Disclosure Stmt	
8021 40	8021 40	Recording each patent assignment per property (times number of properties)	
1809 770	2809 385	Filing a submission after final rejection (37 CFR 1.129(a))	
1810 770	2810 385	For each additional invention to be examined (37 CFR 1.129(b))	
1801 770	2801 385	Request for Continued Examination (RCE)	
1802 900	1802 900	Request for expedited examination of a design application	

Other fee (specify)

*Reduced by Basic Filing Fee Paid

SUBTOTAL (3) (\$)

SUBMITTED BY

Name (Print/Type)	Cathy A. Kodroff	Registration No. (Attorney/Agent)	33,980	Telephone	215-540-9200
Signature	<i>Cathy A. Kodroff</i>	Date	April 28, 2004		

WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.

This collection of information is required by 37 CFR 1.17 and 1.27. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

Express Mail No. EU531571267US

Customer No. 00270

METHODS OF GENERATING CHIMERIC ADENOVIRUSES AND USES FOR SUCH CHIMERIC ADENOVIRUSES

5 BACKGROUND OF THE INVENTION

The presence of humoral immunity (circulating antibodies) to adenovirus capsid proteins is a barrier to the use of adenovirus vectors for gene therapy. The prototype adenovirus vectors that have been developed for gene therapy are based on subgroup C adenoviruses such as that of serotype 5. The prevalence of neutralizing
10 antibodies against subgroup C adenoviruses is generally high in human populations as a result of frequent exposure to these pathogens. This fact is likely to greatly limit the effectiveness of gene therapy vectors based on serotypes such as Ad5.

Analysis of the nature of the protective antibodies against adenoviruses has indicated that the most important target is the major capsid protein, hexon [Wolfhart
15 (1988) *J. Virol* 62, 2321; Gall *et al.* (1996) *J. Virol.* 70, 2116]. Several efforts have been made to engineer the hexon so as to evade the anti-hexon antibodies by making chimeric adenoviruses harboring hexons from other serotypes [Roy *et al.* (1998) *J. Virol.* 72, 6875; US Patent No 5922315; Gall *et al.* (1998) *J. Virol.* 72, 10260; Youil
et al. (2002) *Hum. Gene Ther.* 13, 311; Wu *et al.* (2002) *J. Virol.* 76, 12775].
20 However, this has been largely unsuccessful when exchanges among distant serotypes are attempted.

Alternatively, investigators have proposed using adenovirus vectors that rarely cause human infections or using adenoviruses from non-human sources. However, the lack of a practical manner in which to produce large numbers of such
25 vectors has proved to be a hindrance to developing such vectors.

SUMMARY OF THE INVENTION

The present invention provides a method of modifying adenoviruses having capsids,
30 and particularly, including hexons, from serotypes which are not well adapted for growth in cells useful for adenoviral virion production. The method is useful for

production of scalable amounts of adenoviruses. The modified, or chimeric, adenoviruses are useful for a variety of purposes which are described herein.

The invention further provides novel, isolated, adenovirus SA18 nucleic acid and amino acid sequences, vectors containing same, cell lines containing such SA18 sequences and/or vectors, and uses thereof.

Other aspects and advantages of the present invention will be readily apparent from the following detailed description of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 provides the map of the genome of the simian adenovirus generated by shotgun cloning as described in the examples below.

Fig. 2 provides the map of the recombinant Adhu5-SV25 chimeric virus. The portion of the genome replaced by DNA from Pan7 is indicated.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides chimeric adenoviruses composed of the left terminal end and right terminal end of an adenovirus which can be cultured in the selected host cell, and the internal regions encoding, at a minimum, the capsid proteins of another adenovirus serotype. This invention is particularly advantageous for generating adenoviruses having serotypes which are difficult to culture in a desired cell type. The invention thus permits generation of chimeric adenoviruses of varying serotypes.

In the embodiments illustrated herein, chimeric adenoviruses have been constructed where most structural proteins, and not merely the hexon or fiber, are derived from an adenovirus of an unrelated serotype, thereby preserving the majority of the protein-protein interactions that are involved in capsid assembly. Most of the early genes such as those encoded by the adenovirus E1 and E4 regions that are responsible for transcription regulation and regulation of the host cell cycle, are retained from a different serotype that is known to result in high titer virus generation in the commonly used cell types, such as HEK 293 which supplies the Ad5 E1 proteins *in trans*.

In another embodiment, the invention provides novel nucleic acid and amino acid sequences from Ad SA18, which was originally isolated from vervet monkey

[ATCC VR-943]. The present invention further provides novel adenovirus vectors and packaging cell lines to produce those vectors for use in the *in vitro* production of recombinant proteins or fragments or other reagents. The invention further provides compositions for use in delivering a heterologous molecule for therapeutic or vaccine purposes. Such therapeutic or vaccine compositions contain the adenoviral vectors carrying an inserted heterologous molecule. In addition, novel sequences of the invention are useful in providing the essential helper functions required for production of recombinant adeno-associated viral (AAV) vectors. Thus, the invention provides helper constructs, methods and cell lines which use these sequences in such production methods.

The term "substantial homology" or "substantial similarity," when referring to a nucleic acid or fragment thereof, indicates that, when optimally aligned with appropriate nucleotide insertions or deletions with another nucleic acid (or its complementary strand), there is nucleotide sequence identity in at least about 95 to 99% of the aligned sequences.

The term "substantial homology" or "substantial similarity," when referring to amino acids or fragments thereof, indicates that, when optimally aligned with appropriate amino acid insertions or deletions with another amino acid (or its complementary strand), there is amino acid sequence identity in at least about 95 to 99% of the aligned sequences. Preferably, the homology is over full-length sequence, or a protein thereof, or a fragment thereof which is at least 8 amino acids, or more desirably, at least 15 amino acids in length. Examples of suitable fragments are described herein.

The term "percent sequence identity" or "identical" in the context of nucleic acid sequences refers to the residues in the two sequences that are the same when aligned for maximum correspondence. The length of sequence identity comparison may be over the full-length of the genome (e.g., about 36 kbp), the full-length of an open reading frame of a gene, protein, subunit, or enzyme [see, e.g., the tables providing the adenoviral coding regions], or a fragment of at least about 500 to 5000 nucleotides, is desired. However, identity among smaller fragments, e.g., of at least about nine nucleotides, usually at least about 20 to 24 nucleotides, at least about 28 to 32 nucleotides, at least about 36 or more nucleotides, may also be desired. Similarly, "percent sequence identity" may be readily determined for amino acid sequences,

over the full-length of a protein, or a fragment thereof. Suitably, a fragment is at least about 8 amino acids in length, and may be up to about 700 amino acids. Examples of suitable fragments are described herein.

5 Identity is readily determined using such algorithms and computer programs as are defined herein at default settings. Preferably, such identity is over the full length of the protein, enzyme, subunit, or over a fragment of at least about 8 amino acids in length. However, identity may be based upon shorter regions, where suited to the use to which the identical gene product is being put.

10 As described herein, alignments are performed using any of a variety of publicly or commercially available Multiple Sequence Alignment Programs, such as "Clustal W", accessible through Web Servers on the internet. Alternatively, Vector NTI utilities are also used. There are also a number of algorithms known in the art that can be used to measure nucleotide sequence identity, including those contained in the programs described above. As another example, polynucleotide sequences can be
15 compared using Fasta, a program in GCG Version 6.1. Fasta provides alignments and percent sequence identity of the regions of the best overlap between the query and search sequences. For instance, percent sequence identity between nucleic acid sequences can be determined using Fasta with its default parameters (a word size of 6 and the NOPAM factor for the scoring matrix) as provided in GCG Version 6.1,
20 herein incorporated by reference. Similarly programs are available for performing amino acid alignments. Generally, these programs are used at default settings, although one of skill in the art can alter these settings as needed. Alternatively, one of skill in the art can utilize another algorithm or computer program that provides at least the level of identity or alignment as that provided by the referenced algorithms and
25 programs.

As used throughout this specification and the claims, the term "comprise" and its variants including, "comprises", "comprising", among other variants, is inclusive of other components, elements, integers, steps and the like. The term "consists of" or "consisting of" are exclusive of other components, elements, integers, steps and the
30 like.

Except where otherwise specified, the term "vector" includes any genetic element known in the art which will deliver a target molecule to a cell, including, naked DNA, a plasmid, phage, transposon, cosmids, episomes, viruses, etc.

By "minigene" is meant the combination of a selected heterologous gene and the other regulatory elements necessary to drive translation, transcription and/or expression of the gene product in a host cell.

As used herein, the term "transcomplement" refers to when a gene (gene product) of one adenovirus serotype supplies an adenovirus serotype lacking this gene (gene product) from another serotype with the missing function. For example, human adenovirus serotype 5 E1a and E1b functions are known to transcomplement E1-deleted chimpanzee adenovirus Pan 9. Similarly, the inventors have found that human Ad5 E1 transcomplements E1-deleted chimpanzee adenovirus serotypes Pan5, Pan6, Pan7, and simian adenovirus serotypes SV1, SV25 and SV39. Other examples of transcomplementing serotypes include human Ad5 and human Ad2, Ad3, Ad4, Ad5, Ad7, and Ad12.

The term "functionally deleted" or "functional deletion" means that a sufficient amount of the gene region is removed or otherwise damaged, e.g., by mutation or modification, so that the gene region is no longer capable of producing functional products of gene expression. If desired, the entire gene region may be removed. Other suitable sites for gene disruption or deletion are discussed elsewhere in the application.

The term "functional" refers to a product (e.g., a protein or peptide) which performs its native function, although not necessarily at the same level as the native product. The term "functional" may also refer to a gene which encodes and from which a desired product can be expressed.

I. Chimeric Adenoviral Vectors

The compositions of this invention include chimeric adenoviral vectors that deliver a heterologous molecule to cells. For delivery of such a heterologous molecule, the vector can be a plasmid or, preferably, a chimeric adenovirus. The chimeric adenoviruses of the invention include adenovirus DNA from at least two source serotypes, a "donating serotype" and a "parental adenovirus" as described in more detail herein, and a minigene.

Because the adenoviral genome contains open reading frames on both strands, in many instances reference is made herein to 5' and 3' ends of the various regions to avoid confusion between specific open reading frames and gene regions. Thus, when

reference is made herein to the "left" and "right" end of the adenoviral genome, this reference is to the ends of the approximately 36 kb adenoviral genome when depicted in schematic form as is conventional in the art [see, e.g., Horwitz, "Adenoviridae and Their Replication", in *VIROLOGY*, 2d ed., pp. 1679-1721 (1990)]. Thus, as used

5 herein, the "left terminal end" of the adenoviral genome refers to portion of the adenoviral genome which, when the genome is depicted schematically in linear form, is located at the extreme left end of the schematic. Typically, the left end refers to be portion of the genome beginning at map unit 0 and extending to the right to include at least the 5' inverted terminal repeats (ITRs), and excludes the internal regions of the

10 genome encoding the structural genes. As used herein, the "right terminal end" of the adenoviral genome refers to portion of the adenoviral genome which, when the genome is depicted schematically in linear form, is located at the extreme right end of the schematic. Typically, the right end of the adenoviral genome refers to be portion of the genome ending at map unit 36 and extending to the left to include at least the 3'

15 ITRs, and excludes the internal regions of the genome encoding the structural genes.

A. Adenovirus Regulatory Sequences

1. Serotype

The selection of the adenovirus serotype donating its left terminal end and right terminal end can be readily made by one of skill in the art from

20 among serotypes which can readily be cultured in the desired cell line. Among other factors which may be considered in selecting the serotype of the donating serotype is compatibility with the adenovirus serotype which will be supplying the internal regions at the location at which their sequences are hybridized.

Suitable adenoviruses for donating their left and right termini

25 are available from the American Type Culture Collection, Manassas, Virginia, US (ATCC), a variety of academic and commercial sources, or the desired regions of the donating adenoviruses may be synthesized using known techniques with reference to sequences published in the literature or available from databases (e.g., GenBank, etc.). Examples of suitable donating adenoviruses include, without limitation, human

30 adenovirus serotypes 2, 3, 4, 5, 7, and 12, and further including any of the presently identified human types [see, e.g., Horwitz, "Adenoviridae and Their Replication", in *VIROLOGY*, 2d ed., pp. 1679-1721 (1990)] which can be cultured in the desired cell. Similarly adenoviruses known to infect non-human primates (e.g., chimpanzees,

rhinus, macaque, and other simian species) or other non-human mammals and which grow in the desired cell can be employed in the vector constructs of this invention. Such serotypes include, without limitation, chimpanzee adenoviruses Pan 5 [VR-591], Pan6 [VR-592], Pan7 [VR-593], and C68 (Pan9), described in US Patent No. 6,083,716; and simian adenoviruses including, without limitation SV1 [VR-195]; SV25 [SV-201]; SV35; SV15; SV-34; SV-36; SV-37, and baboon adenovirus [VR-275], among others. In the following examples, the human 293 cells and adenovirus type 5 (Ad5), Pan9, and Ad40 are used for convenience. However, one of skill in the art will understand that other cell lines and/or comparable regions derived from other adenoviral strains may be readily selected and used in the present invention in the place of (or in combination with) these serotypes.

2. Sequences

The minimum sequences which must be supplied by the adenovirus donating its left terminal end and its right terminal end include the 5' cis-elements and the 3' cis-elements necessary for replication and packaging. Typically, the 5' cis-elements necessary for packaging and replication include the 5' inverted terminal repeat (ITR) sequences (which functions as origins of replication) and the native 5' packaging enhancer domains (that contain sequences necessary for packaging linear Ad genomes and enhancer elements for the E1 promoter). The right end of the adenoviral genome includes the 3' cis-elements (including the ITRs) necessary for packaging and encapsidation. Desirably, the adenovirus serotype donating its left and right termini and/or an adenovirus serotype which transcomplements the serotype of the donating adenovirus, further provides the functions of the necessary adenovirus early genes, including E1 (E1a and E1b), E2 (E2a and E2b), and E4 (including at least the ORF6 region). E3 is not essential and may be deleted as desired, e.g., for insertion of a transgene in this region or to provide space for a transgene inserted in another region (typically for packaging it is desirable for the total adenoviral genome to be under 36 kb).

In certain embodiments, the necessary adenovirus early genes are contained in the chimeric construct of the invention. In other embodiment, one or more of the necessary adenovirus early genes can be provided by the packaging host cell or *in trans*.

In general, the chimeric adenovirus of the invention contains regulatory sequences from the donating adenovirus serotype, or a transcomplementing serotype, to provide the chimeric adenovirus with compatible regulatory proteins.

Optionally, one or more of the necessary adenoviral structural genes is provided by the adenovirus donating its left terminal and its right terminal end.

In certain embodiments, the chimeric adenovirus further contains one or more functional adenovirus genes, including, the Endoprotease open reading frame, DNA binding protein, 100 kDa scaffolding protein, 33 kDa protein, protein VIII, pTP, 52/55 kDa protein, protein VII, Mu and/or protein VI from the adenovirus serotype donating its left and right termini. Where all of these genes are derived from the adenovirus serotype donating the 5' and 3' ITRs, a "pseudotyped" chimeric is formed. Optionally, one or more of the genes can be hybrids formed from the fusion of the donating adenovirus serotype and the parental adenovirus serotype providing the capsid proteins (e.g., without limitation, polymerase, terminal protein, IIIa protein). Suitably, these genes express functional proteins which permit packaging of the adenovirus genes into the capsid. Alternatively, one or more of these proteins (whether hybrid or non-hybrid) can be functionally deleted in the chimeric adenovirus. Where desired, any necessary proteins functionally deleted in the chimeric adenovirus can be expressed in *trans* in the packaging cell.

B. Parental Adenovirus Structural Proteins

1. Serotypes

This invention is particularly well adapted for use in generating chimeric adenoviruses in which the capsid proteins are from a parental adenovirus which does not efficiently grow in a desirable host cell. The selection of the parental adenovirus serotype providing the internal regions can be readily made by one of skill in the art based on the information provided herein.

A variety of suitable adenoviruses can serve as a parental adenovirus supplying the regions encoding the structural (i.e., capsid proteins). Many of such adenoviruses can be obtained from the same sources as described above for the donating adenovirus serotypes. Examples of suitable parental adenovirus serotypes includes, without limitation, human adenovirus serotype 40, among others [see, e.g., Horwitz, "Adenoviridae and Their Replication", in *VIROLOGY*, 2d ed., pp. 1679-1721 (1990)], and adenoviruses known to infect non-human primates (e.g.,

chimpanzees, rhesus, macaque, and other simian species) or other non-human mammals, including, without limitation, chimpanzee adenovirus C1, described in US Patent No. 6,083,716; simian adenoviruses, and baboon adenoviruses, among others. In addition, the parental adenovirus supplying the internal regions may be from a non-naturally occurring adenovirus serotype, such as may be generated using a variety of techniques known to those of skill in the art.

In one embodiment illustrated herein, a chimeric virus that was constructed was that between the chimpanzee adenoviruses Pan-5 and C1 exhibited a higher titer in human 293 cells than the wild-type parental virus. However, the invention is not limited to the use of these chimpanzee adenoviruses, or to the combination of simian-simian, human-human, or simian-human chimeric adenoviruses. For example, it may be desirable to utilize bovine or canine adenoviruses, or other non-human mammalian adenoviruses which do not naturally infect and/or replicate in human cells.

In the following examples, the human adenovirus type 40 (Ad40) and the chimpanzee adenovirus C1, simian Pan 5 and Ad40, and Pan 5 and simian adenovirus SA18, are used. However, one of skill in the art will understand that other adenoviral serotypes may be readily selected and used in the present invention in the place of (or in combination with) these serotypes.

2. Sequences

The parental adenovirus provides to the chimeric construct of the invention its internal regions which includes structural proteins necessary for generating a capsid having the desired characteristics of the parental adenovirus. These desired characteristics include, but are not limited to, the ability to infect target cells and delivery a heterologous transgene, the ability to elude neutralizing antibodies directed to another adenovirus serotype (i.e., avoiding clearance due to cross-reactivity), and/or the ability to infect cells in the absence of an immune response to the chimeric adenovirus. The advantages of such characteristics may be most readily apparent in a regimen which involves repeat delivery of adenoviral vectors. The left and right termini of the parent adenovirus, including at least the 5' ITRs, the E1 region, the E4 region and the 3' ITRs are non-functional and, preferably, completely absent. Optionally, all adenovirus regulatory proteins from this parental adenovirus are non-functional and only the structural proteins (or selected structural proteins) are retained.

At a minimum, the parental adenovirus provides the adenoviral late region encoding the hexon protein. Suitably, the parental adenovirus further provides the late regions encoding the penton and the fiber. In certain embodiments, all of the functional adenoviral late regions, including L1 (encoding 52/55 Da, IIIa proteins), L2 (encoding penton, VII, V, Mu proteins), L3 (encoding VI, hexon, Endoprotease), L4 (encoding 100 kD, 33 kD, VIII proteins) and L5 (encoding fiber protein) are supplied by the parental adenovirus. Optionally, one or more of these late gene functions, with the exception of those encoding the hexon, penton and fiber proteins, can be functionally deleted. Any necessary structural proteins may be supplied in *trans*.

Thus, in certain embodiments, the chimeric adenovirus further contains one or more functional adenovirus genes, including, the Endoprotease open reading frame, DNA binding protein, 100 kDa scaffolding protein, 33 kDa protein, protein VIII, pTP, 52/55 kDa protein, protein VII, Mu and/or protein VI from the parental adenovirus donating its internal regions. Optionally, one or more of the genes can be hybrids formed from the fusion of the donating adenovirus serotype and the parental adenovirus serotype providing the capsid proteins, as described above.

C. The "Minigene"

Typically, an adenoviral vector of the invention is designed to contain a minigene which may be inserted into the site of a partially deleted, fully deleted (absent), or disrupted adenoviral gene. For example, the minigene may be located in the site of such a functional E1 deletion or functional E3 deletion, or another suitable site.

The methods employed for the selection of the transgene, the cloning and construction of the "minigene" and its insertion into the viral vector are within the skill in the art given the teachings provided herein.

1. The transgene

The transgene is a nucleic acid sequence, heterologous to the vector sequences flanking the transgene, which encodes a polypeptide, protein, or other product, of interest. The nucleic acid coding sequence is operatively linked to regulatory components in a manner which permits transgene transcription, translation, and/or expression in a host cell.

The composition of the transgene sequence will depend upon the use to which the adenoviral vector will be put. For example, the adenoviral vector may be used as a helper virus in production of recombinant adeno-associated viruses or in production of recombinant adenoviruses deleted of essential adenoviral gene functions which are supplied by the adenoviral vector, or for a variety of production uses. Alternatively, the adenoviral vector may be used for diagnostic purposes.

One type of transgene sequence includes a reporter sequence, which upon expression produces a detectable signal. Such reporter sequences include, without limitation, DNA sequences encoding β -lactamase, β -galactosidase (LacZ), alkaline phosphatase, thymidine kinase, green fluorescent protein (GFP), chloramphenicol acetyltransferase (CAT), luciferase, membrane bound proteins including, for example, CD2, CD4, CD8, the influenza hemagglutinin protein, and others well known in the art, to which high affinity antibodies directed thereto exist or can be produced by conventional means, and fusion proteins comprising a membrane bound protein appropriately fused to an antigen tag domain from, among others, hemagglutinin or Myc. These coding sequences, when associated with regulatory elements which drive their expression, provide signals detectable by conventional means, including enzymatic, radiographic, colorimetric, fluorescence or other spectrographic assays, fluorescent activating cell sorting assays and immunological assays, including enzyme linked immunosorbent assay (ELISA), radioimmunoassay (RIA) and immunohistochemistry. For example, where the marker sequence is the LacZ gene, the presence of the vector carrying the signal is detected by assays for beta-galactosidase activity. Where the transgene is GFP or luciferase, the vector carrying the signal may be measured visually by color or light production in a luminometer.

However, desirably, the transgene is a non-marker sequence encoding a product which is useful in biology and medicine, such as proteins, peptides, RNA, enzymes, or catalytic RNAs. Desirable RNA molecules include tRNA, dsRNA, ribosomal RNA, si RNAs, small hairpin RNAs, trans-splicing RNAs, catalytic RNAs, and antisense RNAs. One example of a useful RNA sequence is a sequence which extinguishes expression of a targeted nucleic acid sequence in the treated animal.

The transgene may be used for treatment, e.g., of genetic deficiencies, as a cancer therapeutic or vaccine, for induction of an immune response, and/or for prophylactic vaccine purposes. As used herein, induction of an immune response refers to the ability of a molecule (e.g., a gene product) to induce a T cell and/or a humoral immune response to the molecule. The invention further includes using multiple transgenes, e.g., to correct or ameliorate a condition caused by a multi-subunit protein. In certain situations, a different transgene may be used to encode each subunit of a protein, or to encode different peptides or proteins. This is desirable when the size of the DNA encoding the protein subunit is large, e.g., for an immunoglobulin, the platelet-derived growth factor, or a dystrophin protein. In order for the cell to produce the multi-subunit protein, a cell is infected with the recombinant virus containing each of the different subunits. Alternatively, different subunits of a protein may be encoded by the same transgene. In this case, a single transgene includes the DNA encoding each of the subunits, with the DNA for each subunit separated by an internal ribozyme entry site (IRES). This is desirable when the size of the DNA encoding each of the subunits is small, e.g., the total size of the DNA encoding the subunits and the IRES is less than five kilobases. As an alternative to an IRES, the DNA may be separated by sequences encoding a 2A peptide, which self-cleaves in a post-translational event. See, e.g., M.L. Donnelly, *et al*, *J. Gen. Virol.*, 78(Pt 1):13-21 (Jan 1997); Furler, S., *et al*, *Gene Ther.*, 8(11):864-873 (June 2001); Klump H., *et al.*, *Gene Ther.*, 8(10):811-817 (May 2001). This 2A peptide is significantly smaller than an IRES, making it well suited for use when space is a limiting factor. However, the selected transgene may encode any biologically active product or other product, e.g., a product desirable for study.

Suitable transgenes may be readily selected by one of skill in the art. The selection of the transgene is not considered to be a limitation of this invention.

2. Vector and Transgene Regulatory Elements

In addition to the major elements identified above for the minigene, the adenoviral vector also includes conventional control elements which are operably linked to the transgene in a manner that permits its transcription, translation and/or expression in a cell transfected with the plasmid vector or infected with the virus produced by the invention. As used herein, "operably linked"

sequences include both expression control sequences that are contiguous with the gene of interest and expression control sequences that act in trans or at a distance to control the gene of interest.

Expression control sequences include appropriate transcription
5 initiation, termination, promoter and enhancer sequences; efficient RNA processing signals such as splicing and polyadenylation (polyA) signals; sequences that stabilize cytoplasmic mRNA; sequences that enhance translation efficiency (i.e., Kozak consensus sequence); sequences that enhance protein stability; and when desired, sequences that enhance secretion of the encoded product. A great number of
10 expression control sequences, including promoters which are native, constitutive, inducible and/or tissue-specific, are known in the art and may be utilized.

Examples of constitutive promoters include, without limitation, the retroviral Rous sarcoma virus (RSV) LTR promoter (optionally with the RSV enhancer), the cytomegalovirus (CMV) promoter (optionally with the CMV
15 enhancer) [see, e.g., Boshart *et al*, *Cell*, 41:521-530 (1985)], the SV40 promoter, the dihydrofolate reductase promoter, the β -actin promoter, the phosphoglycerol kinase (PGK) promoter, and the EF1 α promoter [Invitrogen].

Inducible promoters allow regulation of gene expression and can be regulated by exogenously supplied compounds, environmental factors such as
20 temperature, or the presence of a specific physiological state, e.g., acute phase, a particular differentiation state of the cell, or in replicating cells only. Inducible promoters and inducible systems are available from a variety of commercial sources, including, without limitation, Invitrogen, Clontech and Ariad. Many other systems have been described and can be readily selected by one of skill in the art. For
25 example, inducible promoters include the zinc-inducible sheep metallothionein (MT) promoter and the dexamethasone (Dex)-inducible mouse mammary tumor virus (MMTV) promoter. Other inducible systems include the T7 polymerase promoter system [WO 98/10088]; the ecdysone insect promoter [No *et al*, *Proc. Natl. Acad. Sci. USA*, 93:3346-3351 (1996)], the tetracycline-repressible system [Gossen *et al*,
30 *Proc. Natl. Acad. Sci. USA*, 89:5547-5551 (1992)], the tetracycline-inducible system [Gossen *et al*, *Science*, 268:1766-1769 (1995), see also Harvey *et al*, *Curr. Opin. Chem. Biol.*, 2:512-518 (1998)]. Other systems include the FK506 dimer, VP16 or p65 using castradiol, diphenol murislerone, the RU486-inducible system [Wang *et al*,

Nat. Biotech., 15:239-243 (1997) and Wang *et al.*, *Gene Ther.*, 4:432-441 (1997)] and the rapamycin-inducible system [Magari *et al.*, *J. Clin. Invest.*, 100:2865-2872 (1997)]. The effectiveness of some inducible promoters increases over time. In such cases one can enhance the effectiveness of such systems by inserting multiple

5 repressors in tandem, e.g., TetR linked to a TetR by an IRES. Alternatively, one can wait at least 3 days before screening for the desired function. One can enhance expression of desired proteins by known means to enhance the effectiveness of this system. For example, using the Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element (WPRE).

10 In another embodiment, the native promoter for the transgene will be used. The native promoter may be preferred when it is desired that expression of the transgene should mimic the native expression. The native promoter may be used when expression of the transgene must be regulated temporally or developmentally, or in a tissue-specific manner, or in response to specific

15 transcriptional stimuli. In a further embodiment, other native expression control elements, such as enhancer elements, polyadenylation sites or Kozak consensus sequences may also be used to mimic the native expression.

Another embodiment of the transgene includes a transgene operably linked to a tissue-specific promoter. For instance, if expression in skeletal

20 muscle is desired, a promoter active in muscle should be used. These include the promoters from genes encoding skeletal β -actin, myosin light chain 2A, dystrophin, muscle creatine kinase, as well as synthetic muscle promoters with activities higher than naturally occurring promoters (see Li *et al.*, *Nat. Biotech.*, 17:241-245 (1999)). Examples of promoters that are tissue-specific are known for liver (albumin,

25 Miyatake *et al.*, *J. Virol.*, 71:5124-32 (1997); hepatitis B virus core promoter, Sandig *et al.*, *Gene Ther.*, 3:1002-9 (1996); alpha-fetoprotein (AFP), Arbuthnot *et al.*, *Hum. Gene Ther.*, 7:1503-14 (1996)), bone osteocalcin (Stein *et al.*, *Mol. Biol. Rep.*, 24:185-96 (1997)); bone sialoprotein (Chen *et al.*, *J. Bone Miner. Res.*, 11:654-64 (1996)), lymphocytes (CD2, Hansal *et al.*, *J. Immunol.*, 161:1063-8 (1998);

30 immunoglobulin heavy chain; T cell receptor chain), neuronal such as neuron-specific enolase (NSE) promoter (Andersen *et al.*, *Cell. Mol. Neurobiol.*, 13:503-15 (1993)), neurofilament light-chain gene (Piccioli *et al.*, *Proc. Natl. Acad. Sci. USA*, 88:5611-5

(1991)), and the neuron-specific vgf gene (Piccioli *et al.*, *Neuron*, 15:373-84 (1995)), among others.

Optionally, vectors carrying transgenes encoding therapeutically useful or immunogenic products may also include selectable markers or reporter genes may include sequences encoding geneticin, hygromycin or purimycin resistance, among others. Such selectable reporters or marker genes (preferably located outside the viral genome to be packaged into a viral particle) can be used to signal the presence of the plasmids in bacterial cells, such as ampicillin resistance. Other components of the vector may include an origin of replication. Selection of these and other promoters and vector elements are conventional and many such sequences are available [see, e.g., Sambrook *et al.*, and references cited therein].

These vectors are generated using the techniques and sequences provided herein, in conjunction with techniques known to those of skill in the art. Such techniques include conventional cloning techniques of cDNA such as those described in texts [Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring Harbor, NY], use of overlapping oligonucleotide sequences of the adenovirus genomes, polymerase chain reaction, and any suitable method which provides the desired nucleotide sequence.

20

II. Production of the Recombinant Viral Particle

In one embodiment, the invention provides a method of generating recombinant chimeric adenoviral particles in which the capsid of the chimeric adenovirus is of a serotype incapable of efficient growth in the selected host cell. A vector suitable for production of recombinant chimeric adenoviral particles can be generated by direct cloning. Alternatively, such particles can be generated by homologous recombination between a first vector containing the left end of the chimeric adenoviral genome and a second vector containing the right end of the chimeric adenoviral genome. However, any suitable methodology known to those of skill in the art can be readily utilized to generate a vector suitable to generate a production vector, preferably which contains the entire chimeric adenoviral genome, including a minigene. This production vector is then introduced into a host cell in

30

which the adenoviral capsid protein is assembled and the chimeric adenoviral particle assembled as described.

5 The chimeric adenoviruses of the invention include those in which one or more adenoviral genes are absent, or otherwise rendered non-functional. If any of the missing gene functions are essential to the replication and infectivity of the adenoviral particle, these functions are supplied by a complementation (or transcomplementing) cell line or a helper vector expressing these functions during production of the chimeric adenoviral particle.

10 Examples of chimeric adenoviruses containing such missing adenoviral gene functions include those which are partially or completely deleted in the E1a and/or E1b gene. In such a case, the E1 gene functions can be supplied by the packaging host cell, permitting the chimeric construct to be deleted of E1 gene functions and, if desired, for a transgene to be inserted in this region. Optionally, the E1 gene can be of a serotype which transcomplements the serotype providing the other adenovirus sequences in order to further reduce the possibility of recombination and improve safety. In other embodiments, it is desirable to retain an intact E1a and/or E1b region in the recombinant adenoviruses. Such an intact E1 region may be located in its native location in the adenoviral genome or placed in the site of a deletion in the native adenoviral genome (e.g., in the E3 region).

20 In another example, all or a portion of the adenovirus delayed early gene E3 may be eliminated from the chimeric adenovirus. The function of adenovirus E3 is believed to be irrelevant to the function and production of the recombinant virus particle. Chimeric adenovirus vectors may also be constructed having a deletion of at least the ORF6 region of the E4 gene, and more desirably because of the redundancy in the function of this region, the entire E4 region. Still another vector of this invention contains a deletion in the delayed early gene E2a. Similarly, deletions in the intermediate genes IX and IVa₂ may be useful for some purposes. Optionally, deletions may also be made in selected portions of the late genes L1 through L5, as described above.

30 Other deletions may be made in the other structural or non-structural adenovirus genes. The above-discussed deletions may be used individually, i.e., an adenovirus sequence for use in the present invention may contain deletions in only a single region. Alternatively, deletions of entire genes or portions thereof effective to

destroy their biological activity may be used in any combination. For example, in one exemplary vector, the adenovirus sequence may have deletions of the E1 genes and the E4 gene, or of the E1, E2a and E3 genes, or of the E1 and E3 genes, or of E1, E2a and E4 genes, with or without deletion of E3, and so on. As discussed above, such deletions may be used in combination with other mutations, such as temperature-sensitive mutations, to achieve a desired result.

Examples of suitable transcomplementing serotypes are provided above. The use of transcomplementing serotypes can be particularly advantageous where there is diversity between the Ad sequences in the vector of the invention and the human AdE1 sequences found in currently available packaging cells. In such cases, the use of the current human E1-containing cells prevents the generation of replication-competent adenoviruses during the replication and production process. However, in certain circumstances, it will be desirable to utilize a cell line which expresses the E1 gene products can be utilized for production of an E1-deleted simian adenovirus. Such cell lines have been described. See, e.g., US Patent 6,083,716.

A. Packaging Host Cells

Suitably, the packaging host cell is selected from among cells in which the adenovirus serotype donating the left and right terminal ends of the chimeric genome are capable of efficient growth. The host cells are preferably of mammalian origin, and most preferably are of non-human primate or human origin.

Particularly desirable host cells are selected from among any mammalian species, including, without limitation, cells such as A549 [ATCC Accession No. CCL 185], 911 cells, WEHI, 3T3, 10T1/2, HEK 293 cells or PERC6 (both of which express functional adenoviral E1) [Fallaux, FJ *et al.*, (1998), *Hum Gene Ther*, 9:1909-1917], Saos, C2C12, L cells, HT1080, HepG2, HeLa [ATCC Accession No. CCL 2], KB [CCL 17], Detroit [e.g., Detroit 510, CCL 72] and WI-38 [CCL 75] cells, and primary fibroblast, hepatocyte and myoblast cells derived from mammals including human, monkey, mouse, rat, rabbit, and hamster. These cell lines are all available from the American Type Culture Collection, 10801 University Boulevard, Manassas, Virginia 20110-2209. Other suitable cell lines may be obtained from other sources. The selection of the mammalian species providing the cells is not a limitation of this invention; nor is the type of mammalian cell, i.e., fibroblast, hepatocyte, tumor cell, etc.

As described above, a chimeric adenovirus of the invention can lack one or more functional adenoviral regulatory and/or structural genes which are supplied either by the host cell or in *trans* to effect packaging of the chimeric adenovirus into the viral capsid to generate the viral particle. Thus, the ability of a
5 selected host cell to supply transcomplementing adenoviral sequences may be taken into consideration in selecting a desired host cell.

In one example, the cells are from a stable cell line which expresses adenovirus E1a and E1b functions from a cell line which transcomplements the adenovirus serotype which donates the left and right termini to the chimera of the
10 invention, permitting the chimera to be E1-deleted. Alternatively, where the cell line does not transcomplement the adenovirus donating the termini, E1 functions may be provided by the chimera, or in *trans*.

If desired, one may utilize the sequences provided herein to generate a packaging cell or cell line that expresses, at a minimum, the adenovirus E1 gene from
15 the adenovirus serotype donating the 5' ITR under the transcriptional control of a promoter for expression, or a transcomplementing serotype, in a selected parent cell line. Inducible or constitutive promoters may be employed for this purpose. Examples of such promoters are described in detail elsewhere in this specification. A parent cell is selected for the generation of a novel cell line expressing any desired
20 adenovirus or adenovirus gene, including, e.g., a human Ad5, AdPan5, Pan6, Pan7, SV1, SV25 or SV39 gene. Without limitation, such a parent cell line may be HeLa [ATCC Accession No. CCL 2], A549 [ATCC Accession No. CCL 185], HEK 293, KB [CCL 17], Detroit [e.g., Detroit 510, CCL 72] and WI-38 [CCL 75] cells, among others. Many of these cell lines are all available from the ATCC. Other suitable
25 parent cell lines may be obtained from other sources.

Such E1-expressing cell lines are useful in the generation of chimeric adenovirus E1 deleted vectors. Additionally, or alternatively, the invention provides cell lines that express one or more simian adenoviral gene products, e.g., E1a, E1b, E2a, and/or E4 ORF6, can be constructed using essentially the same procedures for
30 use in the generation of chimeric viral vectors. Such cell lines can be utilized to transcomplement adenovirus vectors deleted in the essential genes that encode those products, or to provide helper functions necessary for packaging of a helper-dependent virus (e.g., adeno-associated virus). The preparation of a host cell

according to this invention involves techniques such as assembly of selected DNA sequences. This assembly may be accomplished utilizing conventional techniques. Such techniques include cDNA and genomic cloning, which are well known and are described in Sambrook et al., cited above, use of overlapping oligonucleotide
5 sequences of the adenovirus genomes, combined with polymerase chain reaction, synthetic methods, and any other suitable methods which provide the desired nucleotide sequence.

In still another alternative, the essential adenoviral gene products are provided in *trans* by the adenoviral vector and/or helper virus. In such an instance, a
10 suitable host cell can be selected from any biological organism, including prokaryotic (e.g., bacterial) cells, and eukaryotic cells, including, insect cells, yeast cells and mammalian cells. Particularly desirable host cells are selected from among any mammalian species, including, without limitation, cells such as A549, WEHI, 3T3, 10T1/2, HEK 293 cells or PERC6 (both of which express functional adenoviral E1)
15 [Fallaux, FJ et al, (1998), Hum Gene Ther, 9:1909-1917], Saos, C2C12, L cells, HT1080, HepG2 and primary fibroblast, hepatocyte and myoblast cells derived from mammals including human, monkey, mouse, rat, rabbit, and hamster. The selection of the mammalian species providing the cells is not a limitation of this invention; nor is the type of mammalian cell, i.e., fibroblast, hepatocyte, tumor cell, etc.

20 B. Helper Vectors

Thus, depending upon the adenovirus gene content of the adenoviral vectors and any adenoviral gene functions expressed from the host cell, a helper vector may be necessary to provide sufficient adenovirus gene sequences necessary to produce an infective recombinant viral particle containing the minigene. See, for
25 example, the techniques described for preparation of a "minimal" human Ad vector in International Patent Application WO96/13597, published May 9, 1996, and incorporated herein by reference. Suitably, these helper vectors may be non-replicating genetic elements, a plasmid, or a virus.

Useful helper vectors contain selected adenovirus gene sequences not
30 present in the adenovirus vector construct and/or not expressed by the packaging cell line in which the vector is transfected. In one embodiment, the helper virus is replication-defective and contains a variety of adenovirus genes in addition to the

sequences described above. Such a helper vector is desirably used in combination with an E1-expressing cell line.

Helper vectors may be formed into poly-cation conjugates as described in Wu *et al*, *J. Biol. Chem.*, 264:16985-16987 (1989); K. J. Fisher and J. M. Wilson, *Biochem. J.*, 299:49 (April 1, 1994). A helper vector may optionally contain a second reporter minigene. A number of such reporter genes are known to the art. The presence of a reporter gene on the helper virus which is different from the transgene on the adenovirus vector allows both the Ad vector and the helper vector to be independently monitored. This second reporter is used to enable separation between the resulting recombinant virus and the helper virus upon purification.

C. Assembly of Viral Particle and Transfection of a Cell Line

Generally, when delivering the vector comprising the minigene by transfection, the vector is delivered in an amount from about 5 µg to about 100 µg DNA, and preferably about 10 to about 50 µg DNA to about 1×10^4 cells to about 1×10^{13} cells, and preferably about 10^5 cells. However, the relative amounts of vector DNA to host cells may be adjusted, taking into consideration such factors as the selected vector, the delivery method and the host cells selected.

Introduction into the host cell of the vector may be achieved by any means known in the art or as disclosed above, including transfection, and infection. One or more of the adenoviral genes may be stably integrated into the genome of the host cell, stably expressed as episomes, or expressed transiently. The gene products may all be expressed transiently, on an episome or stably integrated, or some of the gene products may be expressed stably while others are expressed transiently.

Furthermore, the promoters for each of the adenoviral genes may be selected independently from a constitutive promoter, an inducible promoter or a native adenoviral promoter. The promoters may be regulated by a specific physiological state of the organism or cell (i.e., by the differentiation state or in replicating or quiescent cells) or by exogenously added factors, for example.

Introduction of the molecules (as plasmids or viruses) into the host cell may also be accomplished using techniques known to the skilled artisan and as discussed throughout the specification. In preferred embodiment, standard transfection techniques are used, e.g., CaPO_4 transfection or electroporation.

Assembly of the selected DNA sequences of the adenovirus (as well as the transgene and other vector elements) into various intermediate plasmids, and the use of the plasmids and vectors to produce a recombinant viral particle are all achieved using conventional techniques. Such techniques include direct cloning as described [G. Gao *et al*, *Gene Ther.* 2003 Oct; 10(22):1926-1930; US Patent Publication No. 2003-0092161-A, published May 15, 2003; International Patent Application No. PCT/US03/12405]. Other cloning techniques of cDNA such as those described in texts [Sambrook *et al*, cited above], use of overlapping oligonucleotide sequences of the adenovirus genomes, polymerase chain reaction, and any suitable method which provides the desired nucleotide sequence can be utilized. Standard transfection and co-transfection techniques are employed, *e.g.*, CaPO₄ precipitation techniques. Other conventional methods employed include homologous recombination of the viral genomes, plaquing of viruses in agar overlay, methods of measuring signal generation, and the like.

For example, following the construction and assembly of the desired minigene-containing viral vector, the vector is transfected *in vitro* in the presence of an optional helper vector into the packaging cell line. The functions expressed from the plasmid, packaging cell line and helper virus, if any, permits the adenovirus-transgene sequences in the vector to be replicated and packaged into virion capsids, resulting in the chimeric viral particles. The current method for producing such virus particles is transfection-based. However, the invention is not limited to such methods. The resulting chimeric adenoviruses are useful in transferring a selected transgene to a selected cell.

III. Use of the Chimeric Adenovirus Vectors

The chimeric adenovirus vectors of the invention are useful for gene transfer to a human or veterinary subject (including, non-human primates, non-simian primates, and other mammals) *in vitro*, *ex vivo*, and *in vivo*.

The recombinant adenovirus vectors described herein can be used as expression vectors for the production of the products encoded by the heterologous genes *in vitro*. For example, the recombinant adenoviruses containing a gene inserted into the location of an E1 deletion may be transfected into an E1-expressing cell line as described above. Alternatively, replication-competent adenoviruses may be used

in another selected cell line. The transfected cells are then cultured in the conventional manner, allowing the recombinant adenovirus to express the gene product from the promoter. The gene product may then be recovered from the culture medium by known conventional methods of protein isolation and recovery from culture.

A chimeric adenoviral vector of the invention provides an efficient gene transfer vehicle that can deliver a selected transgene to a selected host cell *in vivo* or *ex vivo* even where the organism has neutralizing antibodies to one or more AAV serotypes. In one embodiment, the rAAV and the cells are mixed *ex vivo*; the infected cells are cultured using conventional methodologies; and the transduced cells are re-infused into the patient. These compositions are particularly well suited to gene delivery for therapeutic purposes and for immunization, including inducing protective immunity.

More commonly, the chimeric adenoviral vectors of the invention will be utilized for delivery of therapeutic or immunogenic molecules, as described below. It will be readily understood for both applications that the recombinant adenoviral vectors of the invention are particularly well suited for use in regimens involving repeat delivery of recombinant adenoviral vectors. Such regimens typically involve delivery of a series of viral vectors in which the viral capsids are alternated. The viral capsids may be changed for each subsequent administration, or after a pre-selected number of administrations of a particular serotype capsid (e.g., one, two, three, four or more). Thus, a regimen may involve delivery of a rAd with a first capsid, delivery with a rAd with a second capsid, and delivery with a third capsid. A variety of other regimens which use the Ad capsids of the invention alone, in combination with one another, or in combination with other Ad serotypes will be apparent to those of skill in the art. Optionally, such a regimen may involve administration of rAd with capsids of non-human primate adenoviruses, human adenoviruses, or artificial (e.g., chimeric) serotypes such as are described herein. Each phase of the regimen may involve administration of a series of injections (or other delivery routes) with a single Ad serotype capsid followed by a series with another Ad serotype capsid. Alternatively, the recombinant Ad vectors of the invention may be utilized in regimens involving other non-adenoviral-mediated delivery systems, including other viral systems, non-viral delivery systems, protein, peptides, and other biologically active molecules.

The following sections will focus on exemplary molecules which may be delivered via the adenoviral vectors of the invention.

A. Ad-Mediated Delivery of Therapeutic Molecules

In one embodiment, the Ad vectors described herein are administered to humans according to published methods for gene therapy. A viral vector of the invention bearing the selected transgene may be administered to a patient, preferably suspended in a biologically compatible solution or pharmaceutically acceptable delivery vehicle. A suitable vehicle includes sterile saline. Other aqueous and non-aqueous isotonic sterile injection solutions and aqueous and non-aqueous sterile suspensions known to be pharmaceutically acceptable carriers and well known to those of skill in the art may be employed for this purpose.

The adenoviral vectors are administered in sufficient amounts to transduce the target cells and to provide sufficient levels of gene transfer and expression to provide a therapeutic benefit without undue adverse or with medically acceptable physiological effects, which can be determined by those skilled in the medical arts. Conventional and pharmaceutically acceptable routes of administration include, but are not limited to, direct delivery to the retina and other intraocular delivery methods, direct delivery to the liver, inhalation, intranasal, intravenous, intramuscular, intratracheal, subcutaneous, intradermal, rectal, oral and other parenteral routes of administration. Routes of administration may be combined, if desired, or adjusted depending upon the transgene or the condition. The route of administration primarily will depend on the nature of the condition being treated.

Dosages of the viral vector will depend primarily on factors such as the condition being treated, the age, weight and health of the patient, and may thus vary among patients. For example, a therapeutically effective adult human or veterinary dosage of the viral vector is generally in the range of from about 100 μ L to about 100 mL of a carrier containing concentrations of from about 1×10^6 to about 1×10^{15} particles, about 1×10^{11} to 1×10^{13} particles, or about 1×10^9 to 1×10^{12} particles. Dosages will range depending upon the size of the animal and the route of administration. For example, a suitable human or veterinary dosage (for about an 80 kg animal) for intramuscular injection is in the range of about 1×10^9 to about 5×10^{12} particles per mL, for a single site. Optionally, multiple sites of administration may be delivered. In another example, a suitable human or veterinary dosage may be

in the range of about 1×10^{11} to about 1×10^{15} particles for an oral formulation. One of skill in the art may adjust these doses, depending the route of administration, and the therapeutic or vaccinal application for which the recombinant vector is employed. The levels of expression of the transgene, or for an immunogen, the level of circulating antibody, can be monitored to determine the frequency of dosage administration. Yet other methods for determining the timing of frequency of administration will be readily apparent to one of skill in the art.

An optional method step involves the co-administration to the patient, either concurrently with, or before or after administration of the viral vector, of a suitable amount of a short acting immune modulator. The selected immune modulator is defined herein as an agent capable of inhibiting the formation of neutralizing antibodies directed against the recombinant vector of this invention or capable of inhibiting cytolytic T lymphocyte (CTL) elimination of the vector. The immune modulator may interfere with the interactions between the T helper subsets (T_{H1} or T_{H2}) and B cells to inhibit neutralizing antibody formation. Alternatively, the immune modulator may inhibit the interaction between T_{H1} cells and CTLs to reduce the occurrence of CTL elimination of the vector. A variety of useful immune modulators and dosages for use of same are disclosed, for example, in Yang *et al.*, *J. Virol.*, 70(9) (Sept 1996); International Patent Application No. WO96/12406, published May 2, 1996; and International Patent Application No. PCT/US96/03035, all incorporated herein by reference. Typically, such immune modulators would be selected when the transgene is a therapeutic which requires repeat delivery.

1. Therapeutic Transgenes

Useful therapeutic products encoded by the transgene include hormones and growth and differentiation factors including, without limitation, insulin, glucagon, growth hormone (GH), parathyroid hormone (PTH), growth hormone releasing factor (GRF), follicle stimulating hormone (FSH), luteinizing hormone (LH), human chorionic gonadotropin (hCG), vascular endothelial growth factor (VEGF), angiopoietins, angiostatin, granulocyte colony stimulating factor (GCSF), erythropoietin (EPO), connective tissue growth factor (CTGF), basic fibroblast growth factor (bFGF), acidic fibroblast growth factor (aFGF), epidermal growth factor (EGF), transforming growth factor α (TGF α), platelet-derived growth factor (PDGF), insulin growth factors I and II (IGF-I and IGF-II), any one of the

transforming growth factor superfamily, including TGF, activins, inhibins, or any of the bone morphogenic proteins (BMP) BMPs 1-15, any one of the heregulin/neuregulin/ARIA/neu differentiation factor (NDF) family of growth factors, nerve growth factor (NGF), brain-derived neurotrophic factor (BDNF), neurotrophins NT-3 and NT-4/5, ciliary neurotrophic factor (CNTF), glial cell line derived neurotrophic factor (GDNF), neurturin, agrin, any one of the family of semaphorins/collapsins, netrin-1 and netrin-2, hepatocyte growth factor (HGF), ephrins, noggin, sonic hedgehog and tyrosine hydroxylase.

Other useful transgene products include proteins that regulate the immune system including, without limitation, cytokines and lymphokines such as thrombopoietin (TPO), interleukins (IL) IL-1 through IL-25 (including, e.g., IL-2, IL-4, IL-12 and IL-18), monocyte chemoattractant protein, leukemia inhibitory factor, granulocyte-macrophage colony stimulating factor, Fas ligand, tumor necrosis factors and, interferons, and, stem cell factor, flk-2/flt3 ligand. Gene products produced by the immune system are also useful in the invention. These include, without limitation, immunoglobulins IgG, IgM, IgA, IgD and IgE, chimeric immunoglobulins, humanized antibodies, single chain antibodies, T cell receptors, chimeric T cell receptors, single chain T cell receptors, class I and class II MHC molecules, as well as engineered immunoglobulins and MHC molecules. Useful gene products also include complement regulatory proteins such as complement regulatory proteins, membrane cofactor protein (MCP), decay accelerating factor (DAF), CR1, CF2 and CD59.

Still other useful gene products include any one of the receptors for the hormones, growth factors, cytokines, lymphokines, regulatory proteins and immune system proteins. The invention encompasses receptors for cholesterol regulation, including the low density lipoprotein (LDL) receptor, high density lipoprotein (HDL) receptor, the very low density lipoprotein (VLDL) receptor, proteins useful in the regulation of lipids, including, e.g., apolipoprotein (apo) A and its isoforms (e.g., ApoA1), apoE and its isoforms including E2, E3 and E4), SRB1, ABC1, and the scavenger receptor. The invention also encompasses gene products such as members of the steroid hormone receptor superfamily including glucocorticoid receptors and estrogen receptors, Vitamin D receptors and other nuclear receptors. In addition, useful gene products include transcription factors such as jun, fos, max, mad, serum response factor (SRF), AP-1, AP2, myb, MyoD and

myogenin, ETS-box containing proteins, TFE3, E2F, ATF1, ATF2, ATF3, ATF4, ZF5, NFAT, CREB, HNF-4, C/EBP, SP1, CCAAT-box binding proteins, interferon regulation factor (IRF-1), Wilms tumor protein, ETS-binding protein, STAT, GATA-box binding proteins, e.g., GATA-3, and the forkhead family of winged helix proteins.

Other useful gene products include, carbamoyl synthetase I, ornithine transcarbamylase, arginosuccinate synthetase, arginosuccinate lyase, arginase, fumarylacetate hydrolase, phenylalanine hydroxylase, alpha-1 antitrypsin, glucose-6-phosphatase, porphobilinogen deaminase, cystathione beta-synthase, branched chain ketoacid decarboxylase, albumin, isovaleryl-coA dehydrogenase, propionyl CoA carboxylase, methyl malonyl CoA mutase, glutaryl CoA dehydrogenase, insulin, beta-glucosidase, pyruvate carboxylate, hepatic phosphorylase, phosphorylase kinase, glycine decarboxylase, H-protein, T-protein, a cystic fibrosis transmembrane regulator (CFTR) sequence, and a dystrophin cDNA sequence. Other useful gene products include those useful for treatment of hemophilia A (e.g., Factor VIII and its variants, including the light chain and heavy chain of the heterodimer, optionally operably linked by a junction), and the B-domain deleted Factor VIII, see US 6,200,560 and 6,221,349], and useful for treatment of hemophilia B (e.g., Factor IX).

Still other useful gene products include non-naturally occurring polypeptides, such as chimeric or hybrid polypeptides having a non-naturally occurring amino acid sequence containing insertions, deletions or amino acid substitutions. For example, single-chain engineered immunoglobulins could be useful in certain immunocompromised patients. Other types of non-naturally occurring gene sequences include antisense molecules and catalytic nucleic acids, such as ribozymes, which could be used to reduce overexpression of a target.

Reduction and/or modulation of expression of a gene are particularly desirable for treatment of hyperproliferative conditions characterized by hyperproliferating cells, as are cancers and psoriasis. Target polypeptides include those polypeptides which are produced exclusively or at higher levels in hyperproliferative cells as compared to normal cells. Target antigens include polypeptides encoded by oncogenes such as myb, myc, fyn, and the translocation gene bcr/abl, ras, src, P53, neu, trk and EGRF. In addition to oncogene products as

target antigens, target polypeptides for anti-cancer treatments and protective regimens include variable regions of antibodies made by B cell lymphomas and variable regions of T cell receptors of T cell lymphomas which, in some embodiments, are also used as target antigens for autoimmune disease. Other tumor-associated polypeptides can be used as target polypeptides such as polypeptides which are found at higher levels in tumor cells including the polypeptide recognized by monoclonal antibody 17-1A and folate binding polypeptides.

Other suitable therapeutic polypeptides and proteins include those which may be useful for treating individuals suffering from autoimmune diseases and disorders by conferring a broad based protective immune response against targets that are associated with autoimmunity including cell receptors and cells which produce self-directed antibodies. T-cell mediated autoimmune diseases include rheumatoid arthritis (RA), multiple sclerosis (MS), Sjögren's syndrome, sarcoidosis, insulin dependent diabetes mellitus (IDDM), autoimmune thyroiditis, reactive arthritis, ankylosing spondylitis, scleroderma, polymyositis, dermatomyositis, psoriasis, vasculitis, Wegener's granulomatosis, Crohn's disease and ulcerative colitis. Each of these diseases is characterized by T cell receptors (TCRs) that bind to endogenous antigens and initiate the inflammatory cascade associated with autoimmune diseases.

The chimeric adenoviral vectors of the invention are particularly well suited for therapeutic regimens in which multiple adenoviral-mediated deliveries of transgenes is desired, e.g., in regimens involving redelivery of the same transgene or in combination regimens involving delivery of other transgenes. Such regimens may involve administration of a chimeric adenoviral vector, followed by re-administration with a vector from the same serotype adenovirus. Particularly desirable regimens involve administration of a chimeric adenoviral vector of the invention, in which the serotype of the viral vector delivered in the first administration differs from the serotype of the viral vector utilized in one or more of the subsequent administrations. For example, a therapeutic regimen involves administration of a chimeric vector and repeat administration with one or more adenoviral vectors of the same or different serotypes. In another example, a therapeutic regimen involves administration of an adenoviral vector followed by repeat administration with a chimeric vector of the invention which differs from the serotype of the first delivered adenoviral vector, and optionally further administration with another vector which is the same or, preferably,

differs from the serotype of the vector in the prior administration steps. These regimens are not limited to delivery of adenoviral vectors constructed using the chimeric serotypes of the invention. Rather, these regimens can readily utilize chimeric or non-chimeric vectors of other adenoviral serotypes, which may be of artificial, human or non-human primate, or other mammalian sources, in combination with one or more of the chimeric vectors of the invention. Examples of such serotypes are discussed elsewhere in this document. Further, these therapeutic regimens may involve either simultaneous or sequential delivery of chimeric adenoviral vectors of the invention in combination with non-adenoviral vectors, non-viral vectors, and/or a variety of other therapeutically useful compounds or molecules. The present invention is not limited to these therapeutic regimens, a variety of which will be readily apparent to one of skill in the art.

B. Ad-Mediated Delivery of Immunogenic Transgenes

The adenoviruses of the invention may also be employed as immunogenic compositions. As used herein, an immunogenic composition is a composition to which a humoral (e.g., antibody) or cellular (e.g., a cytotoxic T cell) response is mounted to a transgene product delivered by the immunogenic composition following delivery to a mammal, and preferably a primate. The present invention provides an Ad that can contain in any of its adenovirus sequence deletions a gene encoding a desired immunogen. Chimeric adenoviruses based on simian or other non-human mammalian primate serotypes are likely to be better suited for use as a live recombinant virus vaccine in different animal species compared to an adenovirus of human origin, but is not limited to such a use. The recombinant adenoviruses can be used as prophylactic or therapeutic vaccines against any pathogen for which the antigen(s) crucial for induction of an immune response and able to limit the spread of the pathogen has been identified and for which the cDNA is available.

Such vaccinal (or other immunogenic) compositions are formulated in a suitable delivery vehicle, as described above. Generally, doses for the immunogenic compositions are in the range defined above for therapeutic compositions. The levels of immunity of the selected gene can be monitored to determine the need, if any, for boosters. Following an assessment of antibody titers in the serum, optional booster immunizations may be desired.

Optionally, a vaccinal composition of the invention may be formulated to contain other components, including, e.g. adjuvants, stabilizers, pH adjusters, preservatives and the like. Such components are well known to those of skill in the vaccine art. Examples of suitable adjuvants include, without limitation, liposomes, alum, monophosphoryl lipid A, and any biologically active factor, such as cytokine, an interleukin, a chemokine, a ligands, and optimally combinations thereof. Certain of these biologically active factors can be expressed in vivo, e.g., via a plasmid or viral vector. For example, such an adjuvant can be administered with a priming DNA vaccine encoding an antigen to enhance the antigen-specific immune response compared with the immune response generated upon priming with a DNA vaccine encoding the antigen only.

The adenoviruses are administered in "an immunogenic amount", that is, an amount of adenovirus that is effective in a route of administration to transfect the desired cells and provide sufficient levels of expression of the selected gene to induce an immune response. Where protective immunity is provided, the recombinant adenoviruses are considered to be vaccine compositions useful in preventing infection and/or recurrent disease.

Alternatively, or in addition, the vectors of the invention may contain a transgene encoding a peptide, polypeptide or protein which induces an immune response to a selected immunogen. The recombinant adenoviruses of this invention are expected to be highly efficacious at inducing cytolytic T cells and antibodies to the inserted heterologous antigenic protein expressed by the vector.

For example, immunogens may be selected from a variety of viral families. Example of desirable viral families against which an immune response would be desirable include, the picornavirus family, which includes the genera rhinoviruses, which are responsible for about 50% of cases of the common cold; the genera enteroviruses, which include polioviruses, coxsackieviruses, echoviruses, and human enteroviruses such as hepatitis A virus; and the genera aphthoviruses, which are responsible for foot and mouth diseases, primarily in non-human animals. Within the picornavirus family of viruses, target antigens include the VP1, VP2, VP3, VP4, and VPG. Another viral family includes the calcivirus family, which encompasses the Norwalk group of viruses, which are an important causative agent of epidemic gastroenteritis. Still another viral family desirable for use in targeting antigens for

inducing immune responses in humans and non-human animals is the togavirus family, which includes the genera alphavirus, which include Sindbis viruses, Ross River virus, and Venezuelan, Eastern & Western Equine encephalitis, and rubivirus, including Rubella virus. The flaviviridae family includes dengue, yellow fever, Japanese encephalitis, St. Louis encephalitis and tick borne encephalitis viruses. Other target antigens may be generated from the Hepatitis C or the coronavirus family, which includes a number of non-human viruses such as infectious bronchitis virus (poultry), porcine transmissible gastroenteric virus (pig), porcine hemagglutinin encephalomyelitis virus (pig), feline infectious peritonitis virus (cats), feline enteric coronavirus (cat), canine coronavirus (dog), and human respiratory coronaviruses, which may cause the common cold and/or non-A, B or C hepatitis. In addition, the human coronaviruses include the putative causative agent of sudden acute respiratory syndrome (SARS). Within the coronavirus family, target antigens include the E1 (also called M or matrix protein), E2 (also called S or Spike protein), E3 (also called HE or hemagglutinin-esterase) glycoprotein (not present in all coronaviruses), or N (nucleocapsid). Still other antigens may be targeted against the rhabdovirus family, which includes the genera vesiculovirus (e.g., Vesicular Stomatitis Virus), and the genus lyssavirus (e.g., rabies). Within the rhabdovirus family, suitable antigens may be derived from the G protein or the N protein. The family filoviridae, which includes hemorrhagic fever viruses such as Marburg and Ebola virus, may be a suitable source of antigens. The paramyxovirus family includes parainfluenza Virus Type 1, parainfluenza Virus Type 3, bovine parainfluenza Virus Type 3, rubulavirus (mumps virus), parainfluenza Virus Type 2, parainfluenza virus Type 4, Newcastle disease virus (chickens), rinderpest, morbillivirus, which includes measles and canine distemper, and pneumovirus, which includes respiratory syncytial virus. The influenza virus is classified within the family orthomyxovirus and is a suitable source of antigen (e.g., the HA protein, the NA protein). The bunyavirus family includes the genera bunyavirus (California encephalitis, La Crosse), phlebovirus (Rift Valley Fever), hantavirus (pneumonia is a hemorrhagic fever virus), hantavirus (Nairobi sheep disease) and various unassigned bunyaviruses. The arenavirus family provides a source of antigens against LCM and Lassa fever virus. The reovirus family includes the genera reovirus, rotavirus (which

causes acute gastroenteritis in children), orbiviruses, and cultivirus (Colorado Tick fever), Lebombo (humans), equine encephalosis, blue tongue.

The retrovirus family includes the sub-family oncorivirinal which encompasses such human and veterinary diseases as feline leukemia virus, HTLV I and HTLV II, lentivirinal (which includes human immunodeficiency virus (HIV), simian immunodeficiency virus (SIV), feline immunodeficiency virus (FIV), equine infectious anemia virus, and spumavirinal). Among the lentiviruses, many suitable antigens have been described and can readily be selected. Examples of suitable HIV and SIV antigens include, without limitation the gag, pol, Vif, Vpx, VPR, Env, Tat, Nef, and Rev proteins, as well as various fragments thereof. For example, suitable fragments of the Env protein may include any of its subunits such as the gp120, gp160, gp41, or smaller fragments thereof, e.g., of at least about 8 amino acids in length. Similarly, fragments of the tat protein may be selected. [See, US Patent 5,891,994 and US Patent 6,193,981.] See, also, the HIV and SIV proteins described in D.H. Barouch *et al*, *J. Virol.*, 75(5):2462-2467 (March 2001), and R.R. Amara, *et al*, *Science*, 292:69-74 (6 April 2001). In another example, the HIV and/or SIV immunogenic proteins or peptides may be used to form fusion proteins or other immunogenic molecules. See, e.g., the HIV-1 Tat and/or Nef fusion proteins and immunization regimens described in WO 01/54719, published August 2, 2001, and WO 99/16884, published April 8, 1999. The invention is not limited to the HIV and/or SIV immunogenic proteins or peptides described herein. In addition, a variety of modifications to these proteins have been described or could readily be made by one of skill in the art. See, e.g., the modified gag protein that is described in US Patent 5,972,596. Further, any desired HIV and/or SIV immunogens may be delivered alone or in combination. Such combinations may include expression from a single vector or from multiple vectors. Optionally, another combination may involve delivery of one or more expressed immunogens with delivery of one or more of the immunogens in protein form. Such combinations are discussed in more detail below.

The papovavirus family includes the sub-family polyomaviruses (BKU and JCU viruses) and the sub-family papillomavirus (associated with cancers or malignant progression of papilloma). The adenovirus family includes viruses (EX, AD7, ARD, O.B.) which cause respiratory disease and/or enteritis. The parvovirus includes family feline parvovirus (feline enteritis), feline panleucopeniavirus, canine

parvovirus, and porcine parvovirus. The herpesvirus family includes the sub-family alphaherpesvirinae, which encompasses the genera simplexvirus (HSV I, HSV II), varicellovirus (pseudorabies, varicella zoster) and the sub-family betaherpesvirinae, which includes the genera cytomegalovirus (HCMV, muromegalovirus) and the sub-family gammaherpesvirinae, which includes the genera lymphocryptovirus, EBV (Burkitt's lymphoma), infectious rhinotracheitis, Marek's disease virus, and rhadinovirus. The poxvirus family includes the sub-family chordopoxvirinae, which encompasses the genera orthopoxvirus (Variola (Smallpox) and Vaccinia (Cowpox)), parapoxvirus, avipoxvirus, capripoxvirus, leporipoxvirus, suipoxvirus, and the sub-family entomopoxvirinae. The hepadnavirus family includes the Hepatitis B virus. One unclassified virus which may be suitable source of antigens is the Hepatitis delta virus. Still other viral sources may include avian infectious bursal disease virus and porcine respiratory and reproductive syndrome virus. The alphavirus family includes equine arteritis virus and various Encephalitis viruses.

The viruses of the present invention may also carry immunogens which are useful to immunize a human or non-human animal against other pathogens including bacteria, fungi, parasitic microorganisms or multicellular parasites which infect human and non-human vertebrates, or from a cancer cell or tumor cell. Examples of bacterial pathogens include pathogenic gram-positive cocci include pneumococci; staphylococci; and streptococci. Pathogenic gram-negative cocci include meningococcus; gonococcus. Pathogenic enteric gram-negative bacilli include enterobacteriaceae; pseudomonas, acinetobacteria and eikenella; melioidosis; salmonella; shigella; haemophilus; moraxella; *H. ducreyi* (which causes chancroid); brucella; *Francisella tularensis* (which causes tularemia); yersinia (pasteurella); streptobacillus moniliformis and spirillum; Gram-positive bacilli include listeria monocytogenes; erysipelas; *Corynebacterium diphtheria* (diphtheria); cholera; *B. anthracis* (anthrax); donovanosis (granuloma inguinale); and bartonellosis. Diseases caused by pathogenic anaerobic bacteria include tetanus; botulism; other clostridia; tuberculosis; leprosy; and other mycobacteria. Pathogenic spirochetal diseases include syphilis; treponematoses: yaws, pinta and endemic syphilis; and leptospirosis. Other infections caused by higher pathogen bacteria and pathogenic fungi include actinomycosis; nocardiosis; cryptococcosis, blastomycosis, histoplasmosis and coccidioidomycosis; candidiasis, aspergillosis, and mucormycosis;

sporotrichosis; paracoccidioidomycosis, petriellidiosis, torulopsosis, mycetoma and chromomycosis; and dermatophytosis. Rickettsial infections include Typhus fever, Rocky Mountain spotted fever, Q fever, and Rickettsialpox. Examples of mycoplasma and chlamydial infections include: mycoplasma pneumoniae; 5 lymphogranuloma venereum; psittacosis; and perinatal chlamydial infections. Pathogenic eukaryotes encompass pathogenic protozoans and helminths and infections produced thereby include: amebiasis; malaria; leishmaniasis; trypanosomiasis; toxoplasmosis; *Pneumocystis carinii*; Trichans; *Toxoplasma gondii*; babesiosis; giardiasis; trichinosis; filariasis; schistosomiasis; nematodes; trematodes 10 or flukes; and cestode (tapeworm) infections.

Many of these organisms and/or toxins produced thereby have been identified by the Centers for Disease Control [(CDC), Department of Health and Human Services, USA], as agents which have potential for use in biological attacks. For example, some of these biological agents, include, *Bacillus anthracis* (anthrax), 15 *Clostridium botulinum* and its toxin (botulism), *Yersinia pestis* (plague), variola major (smallpox), *Francisella tularensis* (tularemia), and viral hemorrhagic fevers [filoviruses (e.g., Ebola, Marburg], and arenaviruses [e.g., Lassa, Machupo]], all of which are currently classified as Category A agents; *Coxiella burnetii* (Q fever); Brucella species (brucellosis), *Burkholderia mallei* (glanders), *Burkholderia* 20 *pseudomallei* (meloidosis), *Ricinus communis* and its toxin (ricin toxin), *Clostridium perfringens* and its toxin (epsilon toxin), Staphylococcus species and their toxins (enterotoxin B), *Chlamydia psittaci* (psittacosis), water safety threats (e.g., *Vibrio cholerae*, *Cryptosporidium parvum*), Typhus fever (*Rickettsia prowazekii*), and viral encephalitis (alphaviruses, e.g., Venezuelan equine encephalitis; eastern equine 25 encephalitis; western equine encephalitis); all of which are currently classified as Category B agents; and Nipah virus and hantaviruses, which are currently classified as Category C agents. In addition, other organisms, which are so classified or differently classified, may be identified and/or used for such a purpose in the future. It will be readily understood that the viral vectors and other constructs described 30 herein are useful to deliver antigens from these organisms, viruses, their toxins or other by-products, which will prevent and/or treat infection or other adverse reactions with these biological agents.

Administration of the vectors of the invention to deliver immunogens against the variable region of the T cells elicit an immune response including CTLs to eliminate those T cells. In RA, several specific variable regions of TCRs which are involved in the disease have been characterized. These TCRs include V-3, V-14, V-17 and V α -17. Thus, delivery of a nucleic acid sequence that encodes at least one of these polypeptides will elicit an immune response that will target T cells involved in RA. In MS, several specific variable regions of TCRs which are involved in the disease have been characterized. These TCRs include V-7 and V α -10. Thus, delivery of a nucleic acid sequence that encodes at least one of these polypeptides will elicit an immune response that will target T cells involved in MS. In scleroderma, several specific variable regions of TCRs which are involved in the disease have been characterized. These TCRs include V-6, V-8, V-14 and V α -16, V α -3C, V α -7, V α -14, V α -15, V α -16, V α -28 and V α -12. Thus, delivery of a chimeric adenovirus that encodes at least one of these polypeptides will elicit an immune response that will target T cells involved in scleroderma.

C. Ad-Mediated Delivery Methods

The therapeutic levels, or levels of immunity, of the selected gene can be monitored to determine the need, if any, for boosters. Following an assessment of CD8+ T cell response, or optionally, antibody titers, in the serum, optional booster immunizations may be desired. Optionally, the adenoviral vectors of the invention may be delivered in a single administration or in various combination regimens, *e.g.*, in combination with a regimen or course of treatment involving other active ingredients or in a prime-boost regimen. A variety of such regimens have been described in the art and may be readily selected.

For example, prime-boost regimens may involve the administration of a DNA (*e.g.*, plasmid) based vector to prime the immune system to a second or further, booster, administration with a traditional antigen, such as a protein or a recombinant virus carrying the sequences encoding such an antigen. See, *e.g.*, WO 00/11140, published March 2, 2000, incorporated by reference. Alternatively, an immunization regimen may involve the administration of a chimeric adenoviral vector of the invention to boost the immune response to a vector (either viral or DNA-based) carrying an antigen, or a protein. In still another alternative, an immunization

regimen involves administration of a protein followed by booster with a vector encoding the antigen.

5 In one embodiment, the invention provides a method of priming and boosting an immune response to a selected antigen by delivering a plasmid DNA vector carrying said antigen, followed by boosting with an adenoviral vector of the invention. In one embodiment, the prime-boost regimen involves the expression of multiproteins from the prime and/or the boost vehicle. See, e.g., R.R. Amara, *Science*, 292:69-74 (6 April 2001) which describes a multiprotein regimen for expression of protein subunits useful for generating an immune response against HIV and SIV. For example, a DNA prime may deliver the Gag, Pol, Vif, VPX and Vpr and Env, Tat, and Rev from a single transcript. Alternatively, the SIV Gag, Pol and HIV-1 Env is delivered in a recombinant adenovirus construct of the invention. Still other regimens are described in WO 99/16884 and WO 01/54719.

15 However, the prime-boost regimens are not limited to immunization for HIV or to delivery of these antigens. For example, priming may involve delivering with a first vector of the invention followed by boosting with a second vector, or with a composition containing the antigen itself in protein form. In one example, the prime-boost regimen can provide a protective immune response to the virus, bacteria or other organism from which the antigen is derived. In another desired embodiment, the prime-boost regimen provides a therapeutic effect that can be measured using convention assays for detection of the presence of the condition for which therapy is being administered.

25 The priming composition may be administered at various sites in the body in a dose dependent manner, which depends on the antigen to which the desired immune response is being targeted. The invention is not limited to the amount or situs of injection(s) or to the pharmaceutical carrier. Rather, the regimen may involve a priming and/or boosting step, each of which may include a single dose or dosage that is administered hourly, daily, weekly or monthly, or yearly. As an example, the mammals may receive one or two doses containing between about 10 μ g to about 50 μ g of plasmid in carrier. A desirable amount of a DNA composition ranges between about 1 μ g to about 10,000 μ g of the DNA vector. Dosages may vary from about 1 μ g to 1000 μ g DNA per kg of subject body weight. The amount or site of delivery is desirably selected based upon the identity and condition of the mammal.

The dosage unit of the vector suitable for delivery of the antigen to the mammal is described herein. The vector is prepared for administration by being suspended or dissolved in a pharmaceutically or physiologically acceptable carrier such as isotonic saline; isotonic salts solution or other formulations that will be apparent to those skilled in such administration. The appropriate carrier will be evident to those skilled in the art and will depend in large part upon the route of administration. The compositions of the invention may be administered to a mammal according to the routes described above, in a sustained release formulation using a biodegradable biocompatible polymer, or by on-site delivery using micelles, gels and liposomes. Optionally, the priming step of this invention also includes administering with the priming composition, a suitable amount of an adjuvant, such as are defined herein.

Preferably, a boosting composition is administered about 2 to about 27 weeks after administering the priming composition to the mammalian subject. The administration of the boosting composition is accomplished using an effective amount of a boosting composition containing or capable of delivering the same antigen as administered by the priming DNA vaccine. The boosting composition may be composed of a recombinant viral vector derived from the same viral source (e.g., adenoviral sequences of the invention) or from another source. Alternatively, the "boosting composition" can be a composition containing the same antigen as encoded in the priming DNA vaccine, but in the form of a protein or peptide, which composition induces an immune response in the host. In another embodiment, the boosting composition contains a DNA sequence encoding the antigen under the control of a regulatory sequence directing its expression in a mammalian cell, e.g., vectors such as well-known bacterial or viral vectors. The primary requirements of the boosting composition are that the antigen of the composition is the same antigen, or a cross-reactive antigen, as that encoded by the priming composition.

In another embodiment, the adenoviral vectors of the invention are also well suited for use in a variety of other immunization and therapeutic regimens. Such regimens may involve delivery of adenoviral vectors of the invention simultaneously or sequentially with Ad vectors of different serotype capsids, regimens in which adenoviral vectors of the invention are delivered simultaneously or sequentially with non-Ad vectors, regimens in which the adenoviral vectors of the invention are delivered simultaneously or sequentially with proteins, peptides, and/or

other biologically useful therapeutic or immunogenic compounds. Such uses will be readily apparent to one of skill in the art.

IV. Simian Adenovirus 18 Sequences

5 The invention provides nucleic acid sequences and amino acid sequences of Ad SA18, which are isolated from the other viral material with which they are associated in nature. These sequences are useful in preparing heterologous molecules containing the nucleic acid sequences and amino acid sequences, and regions or fragments thereof as are described herein, viral vectors which are useful for a variety of purposes, including the constructs and compositions, and such methods as are described herein for the chimeric adenoviruses, including, e.g., in host cells for production of viruses requiring adenoviral helper functions, as delivery vehicles for heterologous molecules such as those described herein. These sequences are also useful in generating the chimeric adenoviruses of the invention.

15 A. Nucleic Acid Sequences

The SA18 nucleic acid sequences of the invention include nucleotides SEQ ID NO: 12, nt 1 to 31967. See, Sequence Listing, which is incorporated by reference herein. The nucleic acid sequences of the invention further encompass the strand which is complementary to the sequences of SEQ ID NO: 12, as well as the RNA and cDNA sequences corresponding to the sequences of these sequences figures and their complementary strands. Further included in this invention are nucleic acid sequences which are greater than 95 to 98%, and more preferably about 99 to 99.9% homologous or identical to the Sequence Listing. Also included in the nucleic acid sequences of the invention are natural variants and engineered modifications of the sequences provided in SEQ ID NO: 12 and their complementary strands. Such modifications include, for example, labels that are known in the art, methylation, and substitution of one or more of the naturally occurring nucleotides with a degenerate nucleotide.

25 The invention further encompasses fragments of the sequences of SA18, their complementary strand, cDNA and RNA complementary thereto. Suitable fragments are at least 15 nucleotides in length, and encompass functional fragments, i.e., fragments which are of biological interest. For example, a functional fragment can express a desired adenoviral product or may be useful in production of

recombinant viral vectors. Such fragments include the gene sequences and fragments listed in the tables below.

5 The following tables provide the transcript regions and open reading frames in the simian adenovirus sequences of the invention. For certain genes, the transcripts and open reading frames (ORFs) are located on the strand complementary to that presented in SEQ ID NO: 12. See, e.g., E2b, E4 and E2a. The calculated molecular weights of the encoded proteins are also shown.

Adenovirus Gene Region	Protein	Ad SA18, SEQ ID NO:12		
		start	End	M.W.
ITR		1	180	
E1a	13S	916	1765	27264
	12S	916	1765	24081
E1b	Small T	1874	2380	19423
	LargeT	2179	3609	52741
	IX	3678	4079	13701
E2b	IVa2	5478	4126	51295
	Polymerase	13745	5229	128392
	PTP	13745	8597	75358
	Agnoprotein	8007	8705	23610
L1	52/55 kD	10788	11945	43416
	IIIa	11966	13699	63999
L2	Penton	13796	15322	57166
	VII	15328	15873	20352
	V	15920	17050	42020
L3	VI	17348	18154	29222
	Hexon	18257	21010	102912
	Endoprotease	21029	21640	23015

Adenovirus Gene Region	Protein	Ad SA18, SEQ ID NO:12		
2a	DBP	23147	21711	53626
L4	100kD	23175	25541	87538
	22 kD homolog	25204	25797	22206
	33 kD homolog	25204	26025	24263
	VIII	26107	26817	25490
E3	Orf #1	26817	27125	11814
L5	Fiber	27192	29015	65455
E4	Orf 6/7	30169	29067	13768
	Orf 6	30169	29303	33832
	Orf 4	30464	30099	14154
	Orf 3	30816	30466	13493
	Orf 2	31205	30813	14698
	Orf 1	31608	31231	14054
ITR		31788	31967	

The SA18 adenoviral nucleic acid sequences are useful as therapeutic
 and immunogenic agents and in construction of a variety of vector systems and host
 5 cells. Such vectors are useful for any of the purposes described above for the
 chimeric adenovirus. Additionally, these SA18 sequences and products may be used
 alone or in combination with other adenoviral sequences or fragments, or in
 combination with elements from other adenoviral or non-adenoviral sequences. The
 adenoviral sequences of the invention are also useful as antisense delivery vectors,
 10 gene therapy vectors, or vaccine vectors, and in methods of using same. Thus, the
 invention further provides nucleic acid molecules, gene delivery vectors, and host
 cells which contain the Ad sequences of the invention.

For example, the invention encompasses a nucleic acid molecule
 containing simian Ad ITR sequences of the invention. In another example, the
 15 invention provides a nucleic acid molecule containing simian Ad sequences of the

invention encoding a desired Ad gene product. Still other nucleic acid molecule constructed using the sequences of the invention will be readily apparent to one of skill in the art, in view of the information provided herein.

5 In one embodiment, the simian Ad gene regions identified herein may be used in a variety of vectors for delivery of a heterologous molecule to a cell. Examples of such molecules and methods of delivery are provided in Section III herein. For example, vectors are generated for expression of an adenoviral capsid protein (or fragment thereof) for purposes of generating a viral vector in a packaging host cell. Such vectors may be designed for expression in trans. Alternatively, such
10 vectors are designed to provide cells which stably contain sequences which express desired adenoviral functions, e.g., one or more of E1a, E1b, the terminal repeat sequences, E2a, E2b, E4, E4ORF6 region.

In addition, the adenoviral gene sequences and fragments thereof are useful for providing the helper functions necessary for production of helper-
15 dependent viruses (e.g., adenoviral vectors deleted of essential functions or adeno-associated viruses (AAV)). For such production methods, the simian adenoviral sequences of the invention are utilized in such a method in a manner similar to those described for the human Ad. However, due to the differences in sequences between the simian adenoviral sequences of the invention and those of human Ad, the use of
20 the sequences of the invention essentially eliminate the possibility of homologous recombination with helper functions in a host cell carrying human Ad E1 functions, e.g., 293 cells, which may produce infectious adenoviral contaminants during rAAV production.

Methods of producing rAAV using adenoviral helper functions have
25 been described at length in the literature with human adenoviral serotypes: See, e.g., US Patent 6,258,595 and the references cited therein. See, also, US Patent 5,871,982; WO 99/14354; WO 99/15685; WO 99/47691. These methods may also be used in production of non-human serotype AAV, including non-human primate AAV serotypes. The simian adenoviral gene sequences of the invention which provide the
30 necessary helper functions (e.g., E1a, E1b, E2a and/or E4 ORF6) can be particularly useful in providing the necessary adenoviral function while minimizing or eliminating the possibility of recombination with any other adenoviruses present in the rAAV-packaging cell which are typically of human origin. Thus, selected genes or open

reading frames of the adenoviral sequences of the invention may be utilized in these rAAV production methods.

Alternatively, recombinant adenoviral simian vectors of the invention may be utilized in these methods. Such recombinant adenoviral simian vectors may include, e.g., a hybrid simian Ad/AAV in which simian Ad sequences flank a rAAV expression cassette composed of, e.g., AAV 3' and/or 5' ITRs and a transgene under the control of regulatory sequences which control its expression. One of skill in the art will recognize that still other simian adenoviral vectors and/or gene sequences of the invention will be useful for production of rAAV and other viruses dependent upon adenoviral helper.

In still another embodiment, nucleic acid molecules are designed for delivery and expression of selected adenoviral gene products in a host cell to achieve a desired physiologic effect. For example, a nucleic acid molecule containing sequences encoding an adenovirus E1a protein of the invention may be delivered to a subject for use as a cancer therapeutic. Optionally, such a molecule is formulated in a lipid-based carrier and preferentially targets cancer cells. Such a formulation may be combined with other cancer therapeutics (e.g., cisplatin, taxol, or the like). Still other uses for the adenoviral sequences provided herein will be readily apparent to one of skill in the art.

In addition, one of skill in the art will readily understand that the Ad sequences of the invention can be readily adapted for use for a variety of viral and non-viral vector systems for in vitro, ex vivo or in vivo delivery of therapeutic and immunogenic molecules, including any of those identified as being deliverable via the chimeric adenoviruses of the invention. For example, the simian Ad genome of the invention can be utilized in a variety of rAd and non-rAd vector systems. Such vectors systems may include, e.g., plasmids, lentiviruses, retroviruses, poxviruses, vaccinia viruses, and adeno-associated viral systems, among others. Selection of these vector systems is not a limitation of the present invention.

The invention further provides molecules useful for production of the simian and simian-derived proteins of the invention. Such molecules which carry polynucleotides including the simian Ad DNA sequences of the invention can be in the form of a vector.

B. Simian Adenoviral Proteins of the Invention

The invention further provides gene products of the above adenoviruses, such as proteins, enzymes, and fragments thereof, which are encoded by the adenoviral nucleic acids of the invention. The invention further encompasses
5 SA18 proteins, enzymes, and fragments thereof, having the amino acid sequences encoded by these nucleic acid sequences which are generated by other methods. Such proteins include those encoded by the open reading frames identified in the tables above, and fragments thereof.

Thus, in one aspect, the invention provides unique simian adenoviral
10 proteins which are substantially pure, i.e., are free of other viral and proteinaceous proteins. Preferably, these proteins are at least 10% homogeneous, more preferably 60% homogeneous, and most preferably 95% homogeneous.

In one embodiment, the invention provides unique simian-derived capsid proteins. As used herein, a simian-derived capsid protein includes any
15 adenoviral capsid protein that contains a SA18 capsid protein or a fragment thereof, as defined above, including, without limitation, chimeric capsid proteins, fusion proteins, artificial capsid proteins, synthetic capsid proteins, and recombinantly capsid proteins, without limitation to means of generating these proteins.

Suitably, these simian-derived capsid proteins contain one or more
20 SA18 regions or fragments thereof (e.g., a hexon, penton, fiber or fragment thereof) in combination with capsid regions or fragments thereof of different adenoviral serotypes, or modified simian capsid proteins or fragments, as described herein. A "modification of a capsid protein associated with altered tropism" as used herein includes an altered capsid protein, i.e, a penton, hexon or fiber protein region, or
25 fragment thereof, such as the knob domain of the fiber region, or a polynucleotide encoding same, such that specificity is altered. The simian-derived capsid may be constructed with one or more of the simian Ad of the invention or another Ad serotypes which may be of human or non-human origin. Such Ad may be obtained from a variety of sources including the ATCC, commercial and academic sources, or
30 the sequences of the Ad may be obtained from GenBank or other suitable sources.

The amino acid sequences of the simian adenoviruses penton proteins of the invention are provided herein. The AdSA18 penton protein is provided in SEQ ID NO: 13. Suitably, any of these penton proteins, or unique fragments thereof, may

be utilized for a variety of purposes. Examples of suitable fragments include the penton having N-terminal and/or C-terminal truncations of about 50, 100, 150, or 200 amino acids, based upon the amino acid numbering provided above. Other suitable fragments include shorter internal, C-terminal, or N-terminal fragments. Further, the penton protein may be modified for a variety of purposes known to those of skill in the art.

The invention further provides the amino acid sequences of the hexon protein of SA18, SEQ ID NO:14. Suitably, this hexon protein, or unique fragments thereof, may be utilized for a variety of purposes. Examples of suitable fragments include the hexon having N-terminal and/or C-terminal truncations of about 50, 100, 150, 200, 300, 400, or 500 amino acids, based upon the amino acid numbering provided above and in SEQ ID NO: 14. Other suitable fragments include shorter internal, C-terminal, or N-terminal fragments. For example, one suitable fragment the loop region (domain) of the hexon protein, designated DE1 and FG1, or a hypervariable region thereof. Such fragments include the regions spanning amino acid residues about 125 to 443; about 138 to 441, or smaller fragments, such as those spanning about residue 138 to residue 163; about 170 to about 176; about 195 to about 203; about 233 to about 246; about 253 to about 264; about 287 to about 297; about 404 to about 430, about 430 to 550, about 545 to 650; of the simian hexon proteins, with reference to SEQ ID NO: 14. Other suitable fragments may be readily identified by one of skill in the art. Further, the hexon protein may be modified for a variety of purposes known to those of skill in the art. Because the hexon protein is the determinant for serotype of an adenovirus, such artificial hexon proteins would result in adenoviruses having artificial serotypes. Other artificial capsid proteins can also be constructed using the chimp Ad penton sequences and/or fiber sequences of the invention and/or fragments thereof.

In one example, it may be desirable to generate an adenovirus having an altered hexon protein utilizing the sequences of a hexon protein of the invention. One suitable method for altering hexon proteins is described in US Patent 5,922,315, which is incorporated by reference. In this method, at least one loop region of the adenovirus hexon is changed with at least one loop region of another adenovirus serotype. Thus, at least one loop region of such an altered adenovirus hexon protein is a simian Ad hexon loop region of the invention. In one embodiment, a loop region

of the SA18 hexon protein is replaced by a loop region from another adenovirus serotype. In another embodiment, the loop region of the SA18 hexon is used to replace a loop region from another adenovirus serotype. Suitable adenovirus serotypes may be readily selected from among human and non-human serotypes, as described herein. SA18 is selected for purposes of illustration only; the other simian Ad hexon proteins of the invention may be similarly altered, or used to alter another Ad hexon. The selection of a suitable serotype is not a limitation of the present invention. Still other uses for the hexon protein sequences of the invention will be readily apparent to those of skill in the art.

The invention further encompasses the fiber proteins of the simian adenoviruses of the invention. The fiber protein of AdSA18 has the amino acid sequence of SEQ ID NO: 15. Suitably, this fiber protein, or unique fragments thereof, may be utilized for a variety of purposes. One suitable fragment is the fiber knob, which spans about amino acids 247 to 425 of SEQ ID NO: 15. Examples of other suitable fragments include the fiber having N-terminal and/or C-terminal truncations of about 50, 100, 150, or 200 amino acids, based upon the amino acid numbering provided above and in SEQ ID NO: 15. Still other suitable fragments include internal fragments. Further, the fiber protein may be modified using a variety of techniques known to those of skill in the art.

The invention further encompasses unique fragments of the proteins of the invention which are at least 8 amino acids in length. However, fragments of other desired lengths can be readily utilized. In addition, the invention encompasses such modifications as may be introduced to enhance yield and/or expression of an SA18 gene product, e.g., construction of a fusion molecule in which all or a fragment of the SA18 gene product is fused (either directly or via a linker) with a fusion partner to enhance. Other suitable modifications include, without limitation, truncation of a coding region (e.g., a protein or enzyme) to eliminate a pre- or pro-protein ordinarily cleaved and to provide the mature protein or enzyme and/or mutation of a coding region to provide a secretable gene product. Still other modifications will be readily apparent to one of skill in the art. The invention further encompasses proteins having at least about 95% to 99% identity to the SA18 proteins provided herein.

As described herein, vectors of the invention containing the adenoviral capsid proteins of the invention are particularly well suited for use in applications in which the neutralizing antibodies diminish the effectiveness of other Ad serotype based vectors, as well as other viral vectors. The rAd vectors of the invention are particularly advantageous in readministration for repeat gene therapy or for boosting immune response (vaccine titers). Examples of such regimens are provided herein.

Under certain circumstances, it may be desirable to use one or more of the SA18 gene products (e.g., a capsid protein or a fragment thereof) to generate an antibody. The term "an antibody," as used herein, refers to an immunoglobulin molecule which is able to specifically bind to an epitope. Thus, the antibodies of the invention bind, preferably specifically and without cross-reactivity, to a SA18 epitope. The antibodies in the present invention exist in a variety of forms including, for example, high affinity polyclonal antibodies, monoclonal antibodies, synthetic antibodies, chimeric antibodies, recombinant antibodies and humanized antibodies. Such antibodies originate from immunoglobulin classes IgG, IgM, IgA, IgD and IgE.

Such antibodies may be generated using any of a number of methods known in the art. Suitable antibodies may be generated by well-known conventional techniques, e.g. Kohler and Milstein and the many known modifications thereof. Similarly desirable high titer antibodies are generated by applying known recombinant techniques to the monoclonal or polyclonal antibodies developed to these antigens [see, e.g., PCT Patent Application No. PCT/GB85/00392; British Patent Application Publication No. GB2188638A; Amit *et al.*, 1986 *Science*, 233:747-753; Queen *et al.*, 1989 *Proc. Nat'l. Acad. Sci. USA*, 86:10029-10033; PCT Patent Application No. PCT/WO9007861; and Riechmann *et al.*, *Nature*, 332:323-327 (1988); Huse *et al.*, 1988a *Science*, 246:1275-1281]. Alternatively, antibodies can be produced by manipulating the complementarity determining regions of animal or human antibodies to the antigen of this invention. See, e.g., E. Mark and Padlin, "Humanization of Monoclonal Antibodies", Chapter 4, The Handbook of Experimental Pharmacology, Vol. 113, The Pharmacology of Monoclonal Antibodies, Springer-Verlag (June, 1994); Harlow *et al.*, 1999, Using Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, NY; Harlow *et al.*, 1989, Antibodies: A Laboratory Manual, Cold Spring Harbor, New York; Houston *et al.*, 1988, *Proc.*

Natl. Acad. Sci. USA 85:5879-5883; and Bird *et al.*, 1988, *Science* 242:423-426.

Further provided by the present invention are anti-idiotypic antibodies (Ab2) and anti-anti-idiotypic antibodies (Ab3). See, e.g., M. Wettendorff *et al.*, "Modulation of anti-tumor immunity by anti-idiotypic antibodies." In *Idiotypic Network and Diseases*, ed. by J. Cerny and J. Hiernaux, 1990 *J. Am. Soc. Microbiol.*, Washington DC: pp. 203-229]. These anti-idiotypic and anti-anti-idiotypic antibodies are produced using techniques well known to those of skill in the art. These antibodies may be used for a variety of purposes, including diagnostic and clinical methods and kits.

Under certain circumstances, it may be desirable to introduce a detectable label or a tag onto a SA18 gene product, antibody or other construct of the invention. As used herein, a detectable label is a molecule which is capable, alone or upon interaction with another molecule, of providing a detectable signal. Most desirably, the label is detectable visually, e.g. by fluorescence, for ready use in immunohistochemical analyses or immunofluorescent microscopy. For example, suitable labels include fluorescein isothiocyanate (FITC), phycoerythrin (PE), allophycocyanin (APC), coriphosphine-O (CPO) or tandem dyes, PE-cyanin-5 (PC5), and PE-Texas Red (ECD). All of these fluorescent dyes are commercially available, and their uses known to the art. Other useful labels include a colloidal gold label. Still other useful labels include radioactive compounds or elements. Additionally, labels include a variety of enzyme systems that operate to reveal a colorimetric signal in an assay, e.g., glucose oxidase (which uses glucose as a substrate) releases peroxide as a product which in the presence of peroxidase and a hydrogen donor such as tetramethyl benzidine (TMB) produces an oxidized TMB that is seen as a blue color. Other examples include horseradish peroxidase (HRP) or alkaline phosphatase (AP), and hexokinase in conjunction with glucose-6-phosphate dehydrogenase which reacts with ATP, glucose, and NAD⁺ to yield, among other products, NADH that is detected as increased absorbance at 340 nm wavelength.

Other label systems that are utilized in the methods of this invention are detectable by other means, e.g., colored latex microparticles [Bangs Laboratories, Indiana] in which a dye is embedded are used in place of enzymes to form conjugates with the target sequences provide a visual signal indicative of the presence of the resulting complex in applicable assays.

Methods for coupling or associating the label with a desired molecule are similarly conventional and known to those of skill in the art. Known methods of label attachment are described [see, for example, Handbook of Fluorescent probes and Research Chemicals, 6th Ed., R. P. M. Haugland, Molecular Probes, Inc., Eugene, OR, 1996; Pierce Catalog and Handbook, Life Science and Analytical Research Products, Pierce Chemical Company, Rockford, IL, 1994/1995]. Thus, selection of the label and coupling methods do not limit this invention.

The sequences, proteins, and fragments of the invention may be produced by any suitable means, including recombinant production, chemical synthesis, or other synthetic means. Suitable production techniques are well known to those of skill in the art. See, e.g., Sambrook et al, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press (Cold Spring Harbor, NY). Alternatively, peptides can also be synthesized by the well known solid phase peptide synthesis methods (Merrifield, *J. Am. Chem. Soc.*, 85:2149 (1962); Stewart and Young, Solid Phase Peptide Synthesis (Freeman, San Francisco, 1969) pp. 27-62). These and other suitable production methods are within the knowledge of those of skill in the art and are not a limitation of the present invention.

In addition, one of skill in the art will readily understand that the Ad sequences of the invention can be readily adapted for use for a variety of viral and non-viral vector systems for *in vitro*, *ex vivo* or *in vivo* delivery of therapeutic and immunogenic molecules. For example, in one embodiment, the simian Ad capsid proteins and other simian adenovirus proteins described herein are used for non-viral, protein-based delivery of genes, proteins, and other desirable diagnostic, therapeutic and immunogenic molecules. In one such embodiment, a protein of the invention is linked, directly or indirectly, to a molecule for targeting to cells with a receptor for adenoviruses. Preferably, a capsid protein such as a hexon, penton, fiber or a fragment thereof having a ligand for a cell surface receptor is selected for such targeting. Suitable molecules for delivery are selected from among the therapeutic molecules described herein and their gene products. A variety of linkers including, lipids, polyLys, and the like may be utilized as linkers. For example, the simian penton protein may be readily utilized for such a purpose by production of a fusion protein using the simian penton sequences in a manner analogous to that described in Medina-Kauwe LK, et al, *Gene Ther.* 2001 May; 8(10):795-803 and Medina-Kauwe

LK, et al, *Gene Ther.* 2001 Dec; 8(23): 1753-1761. Alternatively, the amino acid sequences of simian Ad protein IX may be utilized for targeting vectors to a cell surface receptor, as described in US Patent Appln 20010047081. Suitable ligands include a CD40 antigen, an RGD-containing or polylysine-containing sequence, and the like. Still other simian Ad proteins, including, e.g., the hexon protein and/or the fiber protein, may be used for these and similar purposes.

Still other adenoviral proteins of the invention may be used as alone, or in combination with other adenoviral protein, for a variety of purposes which will be readily apparent to one of skill in the art. In addition, still other uses for the adenoviral proteins of the invention will be readily apparent to one of skill in the art.

The compositions of this invention include vectors that deliver a heterologous molecule to cells, either for therapeutic or vaccine purposes. Such vectors, containing simian adenovirus DNA of SA18 and a minigene, can be constructed using techniques such as those described herein for the chimeric adenoviruses and such techniques as are known in the art. Alternatively, SA19 may be a source for sequences of the chimeric adenoviruses are described herein.

The following examples are illustrative, and are not intended to limit the invention to those illustrated embodiments.

Example 1 - Construction of Pan5/C1 Chimeric Simian Viruses

Five different adenoviruses initially isolated from the chimpanzee, AdC68 [US Patent 6,083,716], AdPan5, AdPan7, AdPan6 and AdC1 [US Patent 6,083,716] have been sequenced. See, International Application No. PCT/US02/33645, filed November 2002 for the sequences of Pan5 [SEQ ID NO:1], Pan7 [SEQ ID NO:3], and Pan6 [SEQ ID NO:2]. This application also provides sequences for SV1, SV25 and SV39 [SEQ ID No. 4, 5, 6, respectively]. Sequence comparison of the capsid protein sequences predicted that AdC1 clearly belonged to a different serological subgroup than the other four chimpanzee derived adenoviruses.

However, attempts to cultivate AdC1 in HEK293 cells revealed it to be fastidious in its growth characteristics (data not shown) and therefore possibly unsuitable for use as a vector using the currently available E1 complementing cell lines. However, because of the obvious sequence dissimilarity of AdC1 capsid

protein sequence from the other chimpanzee derived adenoviruses (as well as the huAd5), chimeric adenovirus vectors were generated with the capsid characteristics of AdC1. In view of the above-mentioned drawbacks associated with only making hexon changes, more extensive replacements were made in the chimera described herein, i.e., construction of chimeras where the replacement went beyond just the hexon, to achieve two goals. The first was to determine whether making extended replacements would allow for the rescue of viruses containing hexons of unrelated serotypes that may not otherwise be amenable to rescue. The second goal was to test whether the growth characteristics of adenovirus vectors such as AdPan5, that have been found in our laboratory to be able to be grown to high titer for the purpose of manufacture, would also be present in the chimeric virus, particularly when the hexon (and other capsid proteins) are derived from a virus such as AdC1 that are difficult to grow to a high yield in cell lines such as HEK293. An added bonus of extending the replacement to include the fiber protein would be to further increase the antigenic dissimilarity to beyond that afforded by a hexon change alone.

As an alternative to obtaining purified virus as source for adenoviral DNA to sequence, we have resorted to cloning restriction fragments of viral DNA obtained from infected cells ("Hirt prep"). The first adenovirus we have sequenced in this way is Simian Adenovirus. EcoRI digestion of the Simian Adenovirus yielded 7 fragments. Shotgun cloning yielded clones of the 5 internal fragments, which were cloned and sequenced. Completion of the sequencing was carried out by walking towards each of the ends of the genome. The map of the genome is shown in Figure 1.

A. *Construction of Two Pan5/C1 Chimeric Plasmids*

The overall approach towards constructing chimeric viruses was to first assemble the complete E1 deleted virus DNA into a single plasmid flanked by recognition sites for the restriction enzyme SmaI, digest the plasmid DNA with SmaI to release the virus DNA ends, and transfect the DNA into HEK293 cells to determine whether viable chimeric adenovirus could be rescued. Two chimeric virus plasmids were constructed, p5C1short and p5C1long.

The plasmid p5C1short harbors an E1 deleted Pan5 virus where an internal 15226 bp segment (18332 – 33557) has been replaced by a functionally analogous 14127 bp (18531 – 32657) from AdC1. This results in the replacement of

the Pan5 proteins hexon, endoprotease, DNA binding protein, 100 kD scaffolding protein, 33 kD protein, protein VIII, and fiber, as well as the entire E3 region, with the homologous segment from AdC1. The ClaI site at the left end of the AdC1 fragment is at the beginning of the hexon gene and the resulting protein is identical to the C1 hexon. The EcoRI site which constitutes the right end of the AdC1 fragment is within the E4 orf 7 part of the AdC1. The right end was ligated to a PCR generated right end fragment from AdPan5 such that the regenerated orf 7-translation product is chimeric between AdPan5 and AdC1.

The plasmid p5C1long harbors an E1 deleted Pan5 virus where an internal 25603 bp segment (7955 – 33557) has been replaced by a functionally analogous 24712 bp (7946 – 32657) from AdC1. This results in the replacement of the AdPan5 pre-terminal protein, 52/55 kD protein, penton base protein, protein VII, Mu, and protein VI with those from AdC1 in addition to those replaced in p5C1short. The AscI site at the left end of the AdC1 fragment is at the beginning of the DNA polymerase gene and results in a chimeric protein where the first 165 amino acids of the AdPan5 DNA polymerase has been replaced by a 167 amino acid segment from AdC1 DNA polymerase. In this N-terminal region, the homology between the AdPan5 and AdC1 DNA polymerase proteins is 81% (72% identity).

The plasmid pDVP5Mlu which contains the left end of AdPan5 was used as the starting plasmid for the chimeric vector construction.

The plasmid pDVP5Mlu was made as follows. A synthetic DNA fragment harboring recognition sites for the restriction enzymes SmaI, MluI, EcoRI and EcoRV respectively was ligated into pBR322 digested with EcoRI and NdeI so as to retain the origin of replication and the beta-lactamase gene. The left end of Pan5 extending to the MluI site (15135 bp) was cloned into this plasmid between the SmaI and MluI sites. The E1 gene was functionally deleted and replaced by a DNA fragment harboring recognition sites for the extremely rare cutter restriction enzyme sites I-CeuI and PI-SceI). The 2904 base pairs of the right end of Pan-5 was PCR amplified using the primers P5L [GCG CAC GCG TCT CTA TCG ATG AAT TCC ATT GGT GAT GGA CAT GC, SEQ ID NO:7] and P5ITR [GCG CAT TTA AAT CAT CAT CAA TAA TAT ACC TCA AAC, SEQ ID NO:8] using Tgo polymerase (Roche). The PCR product was cut with MluI and SmaI, and cloned between MluI and EcoRV of pDVP5Mlu to yield pPan5Mlu+RE. A 3193 bp fragment extending

from the MluI site (15135) to the ClaI (18328) site of Pan5 was then inserted between the same sites of pPan5Mlu+RE to yield pPan5Cla+RE. The 3671 bp ClaI (18531) to EcoRI (22202) fragment of the adenovirus C1 was cloned into pPan5Cla+RE between ClaI (16111) and EcoRI (16116) to yield pPan5C1delRI. The 10452 bp internal
5 EcoRI fragment of the adenovirus C1 (22202 – 32653) was cloned into the EcoRI site of pPan5C1delRI to yield p5C1short. To construct p5C1long, the AdC1 replacement was further extended by replacing the AscI – ClaI 10379 bp fragment of AdPan5 in p5C1short with the AdC1 AscI – ClaI 10591 bp fragment. Finally a green fluorescent protein (GFP) expression cassette was inserted into both p5C1short and p5C1long
10 between the I-CeuI and P1-SceI sites to yield p5C1shortGFP and p5C1longGFP respectively.

B. *Rescue of chimeric Pan5/C1 recombinant vector adenoviruses*

The plasmids p5C1shortGFP and p5C1longGFP were digested with the restriction enzyme Swal and transfected into HEK 293 cells. A typical adenovirus
15 induced cytopathic effect was observed. The rescue of recombinant chimeric adenovirus from the p5C1longGFP transfection was confirmed by collecting the supernatant from the transfection and re-infecting fresh cells which were found to be transduced as determined by GFP expression. Viral DNA prepared from the chimeric recombinant virus was digested with several restriction enzymes and found to have
20 the expected pattern on electrophoresis (data not shown).

The chimeric adenoviral construct with the shorter replacement p5C1short encodes the C1 proteins hexon and fiber as well as the intervening open reading frames for endoprotease, DNA binding protein, 100 kDa scaffolding protein, 33 kDa protein, and protein VIII. (The E3 region is also included within this region
25 but is unlikely to impact on the viability of the chimeric virus). When the replacement was extended to include the additional AdC1 proteins pTP (pre-terminal protein), 52/55 kDa protein, penton base, protein VII, Mu, and protein VI, there was no difficulty in rescuing viable chimeric virus. In this experiment, the chimeric adenovirus construction strategy utilized the presence of AscI and ClaI restriction
30 enzyme sites present on the genes for DNA polymerase and hexon respectively on both AdPan5 and AdC1.

The reasons for the relatively higher yield of the chimeric virus compared to the wild-type AdC1 virus are not clear. In the growth of the 5C1

chimeric virus in 293 cells, the adenoviral early region gene products of E1 and E4 are derived from Ad5 and AdPan5 respectively. The E1 and E4 gene products bind, regulate and de-repress several cellular transcription complexes and coordinate their activity towards viral multiplication. Thus it is possible that the E1 gene products
5 supplied in *trans* from the 293 cells and the E4 gene products from AdPan5 are more optimal in the human 293 cell background than are the equivalent AdC1 gene products. This may also apply to the activity of the major late promoter whose activity is responsible for the transcription of the capsid protein genes. In the
10 chimeric virus, the major late promoter, and the protein IVa2 which transactivates it, are derived from AdPan5. However the E2 gene products required for adenoviral DNA replication pTP and single-stranded DNA – binding protein are derived from AdC1. The adenoviral DNA polymerase, which complexes with pTP, is chimeric in Ad5C1 but mostly AdPan5 derived.

15 Two plasmids have been constructed where the structural proteins derive from the chimpanzee adenovirus Pan 7 and Simian adenovirus 25 respectively, and the flanking sequences are derived from human Ad5 (the commonly used vector strain). The Adhu5-Pan7 chimeric adenovirus has been rescued, demonstrating that the chimeric virus construction method used to derive the Pan 5 – C1 chimeric virus
20 is broadly applicable. Fig. 2 provides the map of the recombinant Adhu5-SV25 chimeric virus. The portion of the genome replaced by DNA from Pan7 is indicated.

Example 2 – C5C1 chimeric vector of invention as a delivery vehicle for immunogenic compositions

25 A Pan 5 (Simian adenovirus 22, a subgroup E adenovirus) - C1 (Simian adenovirus 21, a subgroup B adenovirus) chimeric expressing the Ebola virus (Zaire) glycoprotein (C5C1-CMVGP) was constructed as a model antigen in order to test the efficacy of the vector C5C1-CMVGP as a vaccine; this vector has been compared it to the Adhu5 based vector (H5-CMVGP). Compared to H5-CMVGP, the C5C1-
30 CMVGP vector yielded only a slightly decreased level of GP expression in transduced A549 cells.

Thereafter, GP-specific T cell and B cell responses elicited in B10BR mice vaccinated with either H5-CMVGP or C5C1-CMVGP vectors were compared. The

C5C1-CMVGP vector appeared to induce lower frequencies of gamma interferon producing CD8+ T cells with kinetics slower than the H5-CMVGP vector as determined by intracellular cytokine staining using a H-2k restricted GP-specific peptide as stimulant. The total IgG response to GP, measured by ELISA, was equivalent in serum from mice vaccinated with the C5C1-CMVGP or the H5-CMVGP vectors. However, the C5C1-CMVGP vector induced a more potent Th1 type response while the H5-CMVGP vector stimulated a more balanced Th1/Th2 type response.

10 Example 3 – Generation of Simian Pan5/Human Ad40 Chimeric Adenovirus and Chimpanzee Pan5/Simian SA18 Chimeric Adenovirus

The construction of plasmids designed to rescue chimeric adenoviruses where the outside flanking regions are derived from the chimpanzee adenovirus AdPan5, and the internal region (containing the structural capsid protein genes) are derived from the human adenovirus Ad40 and the simian adenovirus SA18, are described below.

As described for the Pan5-C1 chimeric adenovirus, the overall approach towards constructing chimeric viruses was to first assemble the complete E1 deleted virus DNA into a single plasmid flanked by recognition sites for the restriction enzyme *Swa*I, digest the plasmid DNA with *Swa*I to release the virus DNA ends, and transfect the DNA into HEK293 cells to determine whether viable chimeric adenovirus could be rescued. Two chimeric virus plasmids were constructed, pPan5-40 and pPan5-SA18 corresponding to the two chimeric adenoviruses referred to above. The plasmid pPan5-40 harbors an E1 deleted Pan5 virus where an internal 22975 bp segment (10400 – 33374) has been replaced by a functionally analogous 21603 bp (10043 – 21603) from Ad40. This results in the replacement of the AdPan5 52/55 kD protein, penton base protein, protein VII, Mu, protein VI, hexon, endoprotease, DNA binding protein, 100 kD scaffolding protein, 33 kD protein, protein VIII, and fiber, as well as the entire E3 region, with the homologous segment from Ad40. Similarly, the plasmid pPan5-SA18 harbors an E1 deleted Pan5 virus where an internal 22975 bp segment (10400 – 33374) has been replaced by a functionally analogous 19015 bp (10573 – 29587) from SA18. This results in the replacement of the AdPan5 52/55 kD protein, penton base protein, protein VII, Mu,

protein VI, hexon, endoprotease, DNA binding protein, 100 kD scaffolding protein, 33 kD protein, protein VIII, and fiber, as well as the entire E3 region, with the homologous segment from SA18.

5 The construction of plasmids designed to rescue chimeric adenoviruses where the outside flanking regions are derived from the chimpanzee adenovirus AdPan5, and the internal region (containing the structural capsid protein genes) are derived from the human adenovirus Ad40 and the simian adenovirus SA18, are described below.

A. *Silent mutagenesis of XbaI site:*

10 The plasmid pDVP5Mlu which contains the left end of AdPan5 was used as the starting plasmid for the chimeric vector construction. As a first step the XbaI site (3820) was mutagenized to destroy the recognition site without changing the coding sequence for polymerase. This was done by first sub-cloning the NdeI (812) to HindIII (4931) fragment into the plasmid pNEB193 (New
15 England Biolabs) using the same restriction sites in pNEB193, to yield pNEBp5. A PCR reaction was performed on pNEBp5 using the primers P5XTOP (GATACCTAGGAACGAGGAGGATTTGATATTG, SEQ ID NO:9) and P5XBOT (ATGTACGCCTCCGCGCTCAC, SEQ ID NO:10) to yield a 591 bp product. The PCR product was cleaved with AvrII and BbvCI and ligated into
20 pNEBp5 cut with XbaI and BbvCI to yield the desired mutation in the plasmid pNEBp5mut. The mutated NdeI-HindIII fragment from pNEBp5mut was ligated back into pDVP5Mlu to yield the desired mutated plasmid pDVP5Mlumut.

B. *Insertion of the Pan5 right end comprising the right ITR and the complete E4 region:*

25 The right end of Pan-5 was PCR amplified (P5RE2PCR) using the primers P5E4 [GATCGAATTCCCACTCTGTACCCCATCTCTG, SEQ ID NO:11] and P5ITR [GCG CAT TTA AAT CAT CAT CAA TAA TAT ACC TCA AAC, SEQ ID NO:8] using Tgo polymerase, cut with EcoRI and SmaI, and cloned between EcoRI and EcoRV of pDVP5Mlumut to yield pPan5Mlumut+RE.

30 C. *Insertion of Ad40 or SA18 structural protein sequences:*

In order to construct p5-40, the Ad40 segment from XbaI (10038) to EcoRI (31642) was ligated into pDVP5Mlumut+RE2 between XbaI (8178) and

EcoRI (12924) in two steps: first, the XbaI (30494) to EcoRI (31642) fragment was inserted, followed by the XbaI (10038) – XbaI (30494) fragment.

To construct pPan5-SA18 the XbaI (10568) to EcoRI (29584) fragment from the simian adenovirus SA18 was inserted into pDVP5Mlumut+RE2 between XbaI (8178) and EcoRI (12924).

The minigene encoding for green fluorescent protein was inserted in place of the E1 deletion between the I-CeuI and PI-SceI sites in pPan5-40 and pPan5-SA18 respectively. This plasmids were purified, digested with SwaI and transfected into 293 cells.

EXAMPLE 4 – Generation of Chimeric Pan6/Pan7 Vectors

A panel of GFP expressing vectors were generated. This panel includes vectors that are chimeric between Pan 6 and Pan 7 where the hexon protein or the fiber protein or both hexon and fiber proteins of the Pan 7 vector have been replaced by those from Pan 6. Balb/C mice were immunized with either Pan 6 or Pan 7 and re-administration (by tail vein injection) was attempted 3 weeks later using each of the five GFP expressing vectors (C6-GFP, C7-GFP, and the three chimeric vectors). Three days later the level of liver transduction was estimated qualitatively by examining liver sections for the presence of GFP expression and quantitatively by estimating copies of GFP DNA by Taqman analysis. The data showed that antibodies to both hexon and fiber are important in preventing re-administration of adenoviral vectors.

All publications cited in this specification are incorporated herein by reference. While the invention has been described with reference to a particularly preferred embodiment, it will be appreciated that modifications can be made without departing from the spirit of the invention. Such modifications are intended to fall within the scope of the appended claims.

What is claimed is:

1. A method of efficiently culturing a chimeric adenovirus in a selected host cell, said chimeric adenovirus being from a parental adenovirus strain incapable of efficient growth in said host cell, said method comprising the steps of:

(a) generating a chimeric adenovirus comprising:

(i) adenovirus sequences of the left terminal end and right terminal end of a first adenovirus which grows in a selected host cell type, said left end region comprising the 5' inverted terminal repeat (ITRs), and said right end region comprising the 3' inverted terminal repeat (ITRs); and

(ii) the internal regions from a parental adenovirus which lacks its native 5' and 3' terminal regions, said internal regions comprising the late genes encoding the penton, hexon, and fiber;

wherein the resulting chimeric adenovirus comprises, from 5' to 3', a left terminal region of the first adenovirus, the internal region of the parental adenovirus, and the right terminal region of the first adenovirus; and

b) culturing said chimeric adenovirus in the presence of functional adenovirus E1a, E1b, and E4 ORF6 genes from the first adenovirus or from an adenovirus serotype which transcomplements the first adenovirus, and further in the presence of necessary adenoviral structural genes from the left end of the adenovirus.

2. The method according to claim 1, wherein the internal region of the parental adenovirus further comprises one or more functional adenovirus genes selected from the group consisting of Endoprotease open reading frame, DNA binding protein, 100 kDa scaffolding protein, 33 kDa protein, protein VIII, pTP, 52/55 kDa protein, protein VII, Mu and protein VI.

3. The method according to claim 1, wherein the polymerase, terminal protein and 52/55 kDa protein functions are provided in *trans*.

4. The method according to claim 1, wherein the first adenovirus further comprises the polymerase, terminal protein and 52/55 kDa protein functions.
5. The method according to claim 1, wherein the chimeric adenovirus comprises the adenoviral late genes 1, 2, 3, 4, and 5 of the parental adenovirus.
6. The method according to claim 1, wherein the selected host cell stably contains one or more of the adenovirus E1a, E1b or E4 ORF6 functions.
7. The method according to claim 1, wherein the chimeric adenovirus comprises one or more of the adenovirus E1a, E1b or E4 ORF6 of the first adenovirus.
8. The method according to claim 1, wherein the first adenovirus is of human origin.
9. The method according to claim 1, wherein the first adenovirus is of simian origin.
10. The method according to claim 1, further comprising the step of isolating the chimeric adenovirus.
11. A method for generating a chimeric adenovirus for growth in a selected host cell, said chimeric adenovirus being derived from a parental adenovirus strain incapable of efficient growth in said host cell, said method comprising the step of generating a chimeric adenovirus comprising:
 - 5' and 3' terminal regions of a first adenovirus which grows in a selected host cell type, said 5' terminal regions comprising the 5' inverted terminal repeat (ITRs) and necessary E1 gene functions, and said 3' terminal regions comprising inverted terminal repeat (ITRs) and necessary E4 gene functions; and
 - internal regions from a parental adenovirus which lacks its native 5' and 3' terminal regions, said internal regions comprising the hexon, penton base and fiber;

wherein the resulting chimeric adenovirus comprises, from 5' to 3', the 5' terminal region of the first adenovirus, the internal region of the parental adenovirus, and the 3' terminal regions of the first adenovirus.

12. A chimeric adenovirus produced according to the method of claim 1.

13. A chimeric adenovirus comprising a hexon protein of a selected adenovirus serotype which is incapable of efficient growth in a selected host cell, said modified adenovirus comprising:

(a) adenovirus sequences of the left terminal end of a first adenovirus which grows in a selected host cell type, said left end region comprising the E1a, E1b and 5' inverted terminal repeat (ITRs);

(b) adenovirus sequences of the internal region of the selected adenovirus serotype which is incapable of efficient growth in the selected host cell, said internal region comprising the genes encoding the penton, hexon and fiber of the selected adenovirus;

(c) adenovirus sequences of the right terminal end of the first adenovirus, said right end region comprising the necessary E4 gene functions and the 3' inverted terminal repeat (ITRs),

wherein the resulting chimeric adenovirus comprises adenoviral structural and regulatory proteins necessary for infection and replication.

14. The chimeric adenovirus according to claim 13, wherein the chimeric adenovirus further comprises the IIIa, 52/55kDa and terminal protein (pTP) of the selected adenovirus serotype.

15. The chimeric adenovirus according to claim 13, wherein chimeric adenovirus comprises the polymerase of the first adenovirus.

16. The chimeric adenovirus according to claim 13, wherein the chimeric adenovirus expresses a functional chimeric protein formed from the first adenovirus and the selected adenovirus, said chimeric protein is selected from the group consisting of polymerase, terminal protein, 52/55 kDa protein, and IIIa.

17. The chimeric adenovirus according to claim 13, wherein the chimeric adenovirus comprises the terminal protein, 52/55 kDa, and/or IIIa of the selected adenovirus.

18. A host cell comprising a chimeric adenovirus according to claim 12.

19. The host cell according to claim 18, wherein said host cell is a human cell.

20. An isolated simian adenovirus nucleic acid sequence selected from the group consisting of:

(a) SA18 having the sequence of nucleic acids 1 to 31967 of SEQ ID NO:12 and

(b) a nucleic acid sequence complementary to the sequence of any of (a) to (f).

21. An isolated simian adenovirus serotype nucleic acid sequence selected from one or more of the group consisting of:

(a) 5' inverted terminal repeat (ITR) sequences;

(b) the adenovirus E1a region, or a fragment thereof selected from among the 13S, 12S and 9S regions;

(c) the adenovirus E1b region, or a fragment thereof selected from among the group consisting of the small T, large T, IX, and IVa2 regions;

(d) the E2b region;

(e) the L1 region, or a fragment thereof selected from among the group consisting of the 28.1 kD protein, polymerase, agnoprotein, 52/55 kD protein, and IIIa protein;

(f) the L2 region, or a fragment thereof selected from the group consisting of the penton, VII, VI, and Mu proteins;

(g) the L3 region, or a fragment thereof selected from the group consisting of the VI, hexon, or endoprotease;

(h) the 2a protein;

(i) the L4 region, or a fragment thereof selected from the group consisting of the 100 kD protein, the 33 kD homolog, and VIII;

(j) the E3 region, or a fragment thereof selected from the group consisting of E3 ORF1, E3 ORF2, E3 ORF3, E3 ORF4, E3 ORF5, E3 ORF6, E3 ORF7, E3 ORF8, and E3 ORF9;

(k) the L5 region, or a fragment thereof selected from a fiber protein;

(l) the E4 region, or a fragment thereof selected from the group consisting of E4 ORF7, E4 ORF6, E4 ORF4, E4 ORF3, E4 ORF2, and E4 ORF1;and

(m) the 3' ITR, of any of SA18 SEQ ID NO:12, or a sequence complementary to any of (a) to (m).

22. A simian adenovirus protein encoded by the nucleic acid sequence according to claim 21.

23. A composition comprising a simian adenovirus capsid protein according to claim 22 linked to a heterologous molecule for delivery to a selected host cell.

24. A method for targeting a cell having an adenoviral receptor comprising delivering to a subject a composition according to claim 23.

25. A nucleic acid molecule comprising a heterologous simian adenoviral sequence according to claim 21.

26. The nucleic acid molecule according to claim 25, wherein said simian adenoviral sequence encodes an adenoviral gene product and is operatively linked to regulatory control sequences which direct expression of the adenoviral gene product in a host cells.

27. The nucleic acid molecule according to claim 25, wherein said simian adenoviral sequence comprises the E1a region of SA18 SEQ ID NO:12.

28. A pharmaceutical composition comprising the nucleic acid molecule according to claim 27 and a physiologically compatible carrier.

29. A recombinant adenovirus having a capsid comprising a protein selected from the group consisting of:

- (a) a hexon protein of SA18, SEQ ID NO 13, or a unique fragment thereof;
- (b) a penton protein of SA18, SEQ ID NO: 14, or a unique fragment thereof;
- (c) a fiber protein of SA18, SEQ ID NO: 15, or a unique fragment thereof.

30. The recombinant adenovirus according to claim 29, wherein the capsid is of an artificial serotype.

31. The recombinant adenovirus according to claim 29, wherein said virus further comprises a heterologous gene operatively linked to sequences which direct expression of said gene in a host cell.

32. The recombinant adenovirus according to claim 29, further comprising 5' and 3' adenovirus cis-elements necessary for replication and encapsidation.

33. The recombinant adenovirus according to claim 29, wherein said vector lacks all or a part of the E1 gene.

34. A host cell comprising a heterologous nucleic acid molecule comprising the nucleic acid sequence according to claim 21.

35. The host cell according to claim 34, wherein said host cell is stably transformed with the nucleic acid molecule.

36. The host cell according to claim 34, wherein said host cell expresses one or more adenoviral gene products from said nucleic acid molecule, said adenoviral gene products selected from the group consisting of E1a, E1b, E2a, and E4 ORF6.

37. The host cell according to claim 34, wherein said host cell is stably transformed with a nucleic acid molecule comprising the simian adenovirus inverted terminal repeats.

38. A composition comprising a recombinant virus according to claim 29 in a pharmaceutically acceptable carrier.

39. A method for delivering a heterologous gene to a mammalian cell comprising introducing into said cell an effective amount of the recombinant virus according to claim 29.

40. A method for repeat administration of a heterologous gene to a mammal comprising the steps of:

- (a) introducing into said mammal a first vector which comprises the heterologous gene and
- (b) introducing into said mammal a second vector which comprises the heterologous gene;
wherein at least the first virus or the second vector is a virus according to claim 29 and wherein the first and second recombinant vector are different.

41. A method for producing a selected gene product comprising infecting a mammalian cell with the recombinant virus according to claim 29, culturing said cell under suitable conditions and recovering from said cell culture the expressed gene product.

42. A method for eliciting an immune response in a mammalian host against an infective agent comprising administering to said host an effective amount of the

recombinant adenovirus of claim 29, wherein said heterologous gene encodes an antigen of the infective agent.

43. The method according to claim 42, comprising the step of priming the host with a DNA vaccine comprising the heterologous gene prior to administering the recombinant adenovirus.

ABSTRACT OF THE DISCLOSURE

5 A method for providing an adenovirus from a serotype which does not grow
efficiently in a desired cell line with the ability to grow in that cell line is described.
The method involves replacing the left and right termini of the adenovirus with the
corresponding termini from an adenovirus which grow efficiently in the desired cell
line. At a minimum, the left terminus spans the 5' inverted terminal repeat, the left
terminus spans the E4 region and the 3' inverted terminal repeat. The resulting
10 chimeric adenovirus contains the internal regions spanning the genes encoding the
penton, hexon and fiber from the serotype which does not grow efficiently in the
desired cell. Also provided are vectors constructed from novel simian adenovirus
sequences and proteins, host cells containing same, and uses thereof.

Fig. 1

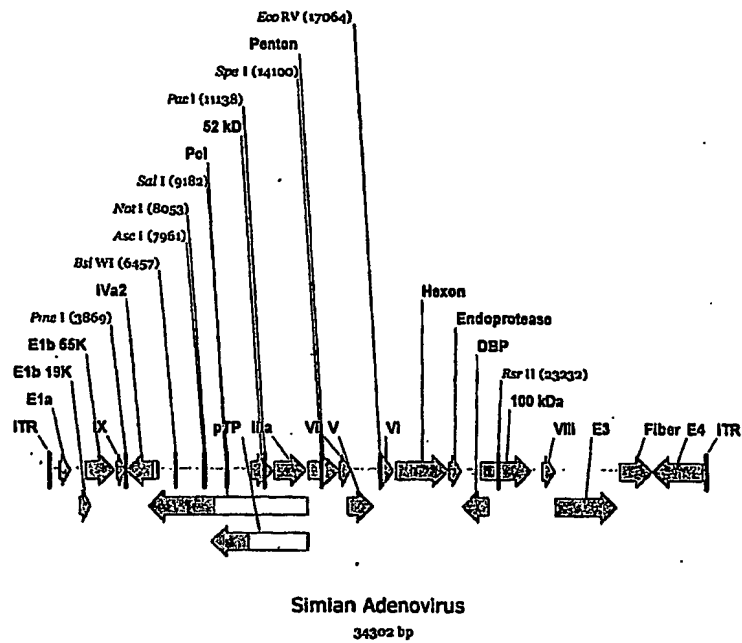
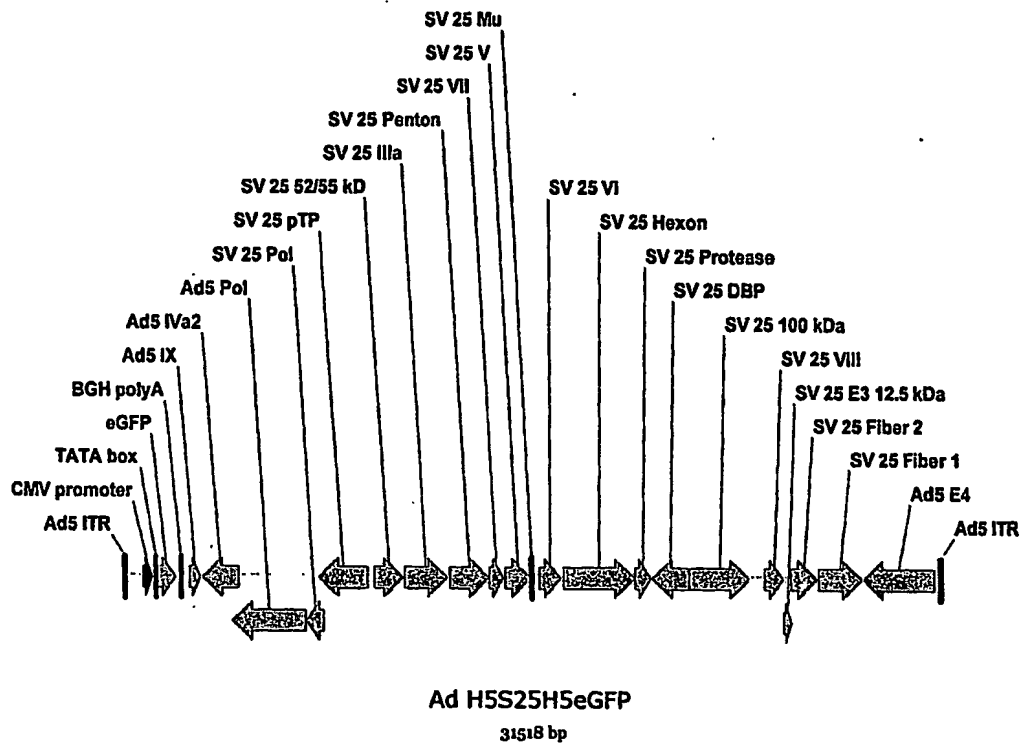


Fig. 2



SEQUENCE LISTING

<110> Wilson, James M.
Roy, Soumitra

<120> Methods of Generating Chimeric Adenoviruses and Uses For Such
Chimeric Adenoviruses

<130> UPN-P3067

<160> 15

<170> PatentIn version 3.2

<210> 1
<211> 36462
<212> DNA
<213> chimpanzee adenovirus serotype Pan5

<400> 1
catcatcaat aatataacctc aaacttttgg tgcgcgttaa tatgcaaatag aggtatttga 60
atttggggat gcggggcggt gattggctgc gggagcgcg accgttaggg gcggggcggg 120
tgacgttttg atgacgtggc cgtgaggcgg agccggtttg caagttctcg tgggaaaagt 180
gacgtcaaac gaggtgtggt ttgaacacgg aaataactcaa ttttcccgcg ctctctgaca 240
ggaaatgagg tgtttctggg cggatgcaag tgaaaacggg ccatpttcgc gcgaaaactg 300
aatgaggaag tgaaaatctg agtaattccg cgtttatggc agggaggagt atttgccgag 360
ggccgagtag actttgaccg attacgtggg ggtttcgatt accgtatttt tcacctaaat 420
ttccgcgtac ggtgtcaaag tccggtgttt ttacgtaggt gtcagctgat cgccagggtgta 480
tttaaacctg cgctctctag tcaagaggcc actcttgagt gccagcgagt agagttttct 540
cctccgcgcc gcgagtcaga tctacacttt gaaagatgag gcacctgaga gacctgcccg 600
gtaatgtttt cctggctact gggaacgaga ttctggaact ggtggtggac gccatgatgg 660
gtgacgaccc tccggagccc cctaccccat ttgaagcgcc ttcgctgtac gatttgtatg 720
atctggaggt ggatgtgccc gagaacgacc ccaacgagga ggcggtgaat gatttgttta 780
gcgatgccgc gctgctggct gccgagcagg ctaatacggg ctctggctca gacagcgatt 840
cctctctcca taccgagaga cccggcagag gtgagaaaaa gatccccgag cttaaagggg 900
aagagctcga cctgcgctgc tatgaggaat gcttgcctcc gagcgatgat gaggaggacg 960
aggaggcgat tcgagctgca gcgaaccagg gagtgaaaac agcgagcgag ggcttttagcc 1020
tggactgtcc tactctgccc ggacacggct gtaagtcttg tgaatttcat cgcataaata 1080
ctggagataa gaatgtgatg tgtgccctgt gctatatgag agcttacaac cattgtgttt 1140

acagtaagtg tgattaactt tagctgggga ggcagagggg gactgggtgc tgactgggtt	1200
atztatgtat atgtttttta tgtgtaggtc cgtctctga cgtagatgag accccacta	1260
cagagtgc atcatcacc ccagaaattg gcgaggaacc gccgaagat attattcata	1320
gaccagttgc agtgagagtc accgggcgta gagcagctgt ggagagtttg gatgacttgc	1380
tacaggggtg ggatgaacct ttggacttgt gtaccggaa acgccccagg cactaagtgc	1440
cacacatgtg tgtttactta aggtgatgtc agtatttata ggggtgtggag tgcaataaaa	1500
tccgtgttga ctttaagtgc gtggtttatg actcaggggt ggggactgtg ggtatataag	1560
cagggtgcaga cctgtgtggt cagttcagag caggactcat ggagatctgg acagtcttgg	1620
aagactttca ccagactaga cagctgctag agaactcacc ggagggagtc tcttacctgt	1680
ggagattctg cttcgggtgg cctctagcta agctagtcta tagggccaag caggattata	1740
aggatcaatt tgaggatatt ttgagagagt gtctgtgtat ttttgactct ctcaacttgg	1800
gccatcagtc tcactttaac cagagtattc tgagagccct tgacttttct actcctggca	1860
gaactaccgc cgcggtagcc ttttttgcct ttatccttga caaatggagt caagaaaccc	1920
atctcagcag ggattaccgt ctggactgct tagcagtagc tttgtggaga acatggaggt	1980
gccagcgct gaatgcaatc tccggctact tgccagtaca gccggtagac acgctgagga	2040
tcttgagtct ccagtcacc caggaacacc aacgcccca gcagccgag caggagcagc	2100
agcaagagga ggaccgagaa gagaacctga gagccgggtct ggacctccg gtggcggagg	2160
aggaggagta gctgacttgt ttcccgagct gcgccgggtg ctgactaggt cttccagtgg	2220
acgggagagg gggattaagc gggagaggca tgaggagact agccacagaa ctgaactgac	2280
tgtcagtctg atgagtcgca ggcgccaga atcggtgtgg tggcatgagg tgcagtcgca	2340
ggggatagat gaggtctcag tgatgcatga gaaatattcc ctagaacaag tcaagacttg	2400
ttgggttgag cccgaggatg attgggaggt agccatcagg aattatgcca agctggctct	2460
gaggccagac aagaagtaca agattaccaa actgattaat atcagaaatt cctgctacat	2520
ttcagggaaat ggggcccagg tggagatcag taccaggag aggggtggcct tcagatgctg	2580
catgatgaat atgtaccgg ggggtgtgg catggaggga gtcaccttta tgaacgcgag	2640
gttcaggggt gatgggtata atgggtgtgt ctttatggcc aacaccaagc tgacagtgca	2700
cggatgctcc ttctttggct tcaataacat gtgcattgag gcctggggca gtgtttcagt	2760
gaggggatgc agtttttcag ccaactggat gggggtcgtg ggcagaacca agagcatgg	2820

gtcagtgaag	aaatgcctgt	tcgagagggtg	ccacctgggg	gtgatgagcg	agggcgaagc	2880
caaagtcaaa	cactgcgcct	ctaccgagac	gggctgcttt	gtactgatca	agggcaatgc	2940
caaagtcaag	cataatatga	tctgtggggc	ctcggatgag	cgcggtacc	agatgctgac	3000
ctgcgcggt	gggaacagcc	atatgctagc	cacggtgcat	gtggcctcgc	accccgcaa	3060
gacatggccc	gagttcgagc	acaacgtcat	gacccgctgc	aatgtgcacc	tggggctccg	3120
ccgaggcatg	ttcatgccct	accagtgcaa	catgcaattt	gtgaagggtgc	tgctggagcc	3180
cgatgccatg	tccagagtga	gcctgacggg	gggtgttgac	atgaatgtgg	agctgtggaa	3240
aattctgaga	tatgatgaat	ccaagaccag	gtgccggggc	tgcaatgcg	gaggcaagca	3300
cgccaggctt	cagcccggtg	gtgtggagg	gacggaggac	ctgcgacccg	atcatttgg	3360
gttgctctgc	aacgggacgg	agttcggtc	cagcggggaa	gaatctgact	agagtgaagta	3420
gtgtttggga	ctgggtggga	gcctgcatga	tgggcagaat	gactaaaatc	tgtgtttttc	3480
tgcgagcag	catgagcgga	agcgctcct	ttgaggagg	ggatttcagc	ccttatctga	3540
cggggcgtct	cccctcctgg	gcgggagtgc	gtcagaatgt	gatgggatcc	acggtggacg	3600
gccggccgt	gcagcccg	aactctcaa	ccctgacct	cgcgacctg	agctcctcgt	3660
ccgtggacgc	agtgccgc	gcagctgctg	cttcgcgc	cagcgccgtg	cgcggaatgg	3720
ccctgggcgc	cggctactac	agctctctgg	tggccaactc	gagttccacc	aataatccc	3780
ccagcctgaa	cgaggagaag	ctgctgctgc	tgatggcca	gctcgaggcc	ctgaccagc	3840
gcctgggcga	gctgaccag	cagggtggctc	agctgcaggc	ggagacgcg	gccgcggtg	3900
ccacggtgaa	aaccaaataa	aaaatgaatc	aataaataa	cggagacggt	tggtgatttt	3960
aacacagagt	cttgaatctt	tatttgattt	ttcgcgcg	gtaggccctg	gaccaccggt	4020
ctcgatcatt	gagcaccgg	tggatctttt	ccaggacctg	gtagagggtg	gcttggtatg	4080
tgaggtagat	gggcatgagc	ccgtcccggg	gggtggaggta	gctccattgc	agggcctcgt	4140
gctcgggggt	gggtgtgtaa	atcaccag	catagcagg	gcgcaggcg	tgggtgctgca	4200
cgatgtcctt	gaggaggaga	ctgatggcca	cgggcagccc	cttggtgtag	gtgttgacga	4260
acctgttgag	ctgggaggga	tgcatgcggg	gggagatgag	atgcatcttg	gcctggatct	4320
tgagattggc	gatgttccc	cccagatccc	gccgggggtt	catgttgtgc	aggaccacca	4380
gcacggtgta	tccggtgcac	ttggggaatt	tgtcatgcaa	cttggaagg	aaggcgtgaa	4440
agaatttgga	gacgcccttg	tgaccgcca	ggttttccat	gcactcatcc	atgatgatgg	4500
cgatgggccc	gtgggcggcg	gcttgggcaa	agacgtttcg	ggggtcggac	acatcgtagt	4560

tgtggtcctg ggtgagctcg tcataggcca ttttaatgaa tttggggcgg aggggtgcccg	4620
actgggggac gaaggtgccc tcatcccg gggcgtagtt gccctcgag atctgcatct	4680
cccaggcctt gagctcggag ggggggatca tgtccacctg cggggcgatg aaaaaaacgg	4740
tttccggggc gggggagatg agctggggccg aaagcagggt cgggagcagc tgggacttgc	4800
cgcagccggt ggggccgtag atgaccccg tgaccggctg cagggtgtag ttgagggaga	4860
gacagctgcc gtcctcgcg aggagggggg ccacctcgtt catcatctcg cgacatgca	4920
tgttctcgcg cagagttcc gccaggaggc gtcgcccc aagcgagagg agctcttgca	4980
gcgaggcgaa gtttttcagc ggcttgagcc cgtcggccat gggcattttg gagagggtct	5040
gttgcaagag ttccagacgg tccagagct cggtagatgt ctctagggca tctcgatcca	5100
gcagacctcc tcgtttcgcg ggttggggcg actgcgggag tagggcacca ggcgatgggc	5160
gtccagcgag gccagggtcc ggtccttcca ggggcgcagg gtccgcgtca gcgtggtctc	5220
cgtcacggtg aaggggtgcg cgcggggctg ggcgcttgcg aggggtgcgt tcaggctcat	5280
ccggctggtc gagaaccgct cccggtcggc gccctgcgcg tcggccagggt agcaattgag	5340
catgagttcg tagttgagcg cctcggccgc gtggcccttg gcgcggagct tacctttgga	5400
agtgtgtccg cagacgggac agaggagggg cttgagggcg tagagcttgg gggcgaggaa	5460
gacggactcg gggcgtagg cgtccgcgcc gcagctggcg cagacggtct cgcactccac	5520
gagccaggtg aggtctggcc ggtcggggc aaaaacgagg tttcctccgt gctttttgat	5580
gcgtttctta cctctggtct ccatgagctc gtgtccccgc tgggtgacaa agaggctgtc	5640
cgtgtccccg tagaccgact ttatgggccc gtctcagagc ggggtgccgc ggtcctcgtc	5700
gtagaggaa cccgcccact ccgagacgaa ggcccgggtc caggccagca cgaaggaggc	5760
cacgtgggag gggtagcggg cgttgtccac cagcgggtcc accttctcca gggtatgcaa	5820
gcacatgtcc ccctcgtcca catccaggaa ggtgattggc ttgtaagtgt aggccacgtg	5880
accgggggtc ccggccgggg ggggtataaaa gggggcgggc ccctgctcgt cctcactgtc	5940
ttccggatcg ctgtccagga gcgccagctg ttggggtagg tattccctct cgaaggcggg	6000
catgacctcg gcactcagggt tgtcagtttc tagaaacgag gaggatttga tattgacggg	6060
gccgttgag acgcctttca tgagccccgc gtccatctgg tcagaaaaga cgatcttttt	6120
gttgtcgagc ttggtggcga aggagccgta gaggcggtg gagagcagct tggcgatgga	6180
gcgcaggtc tgggtttttt ccttgtcggc gcgctccttg gcggcgatgt tgagctgcac	6240

gtactcgcgc	gccacgcact	tccattcggg	gaagacgggtg	gtgagcttgt	cgggcacgat	6300
tctgaccgc	cagccgcggt	tgtgcagggt	gatgagggtcc	acgctgggtg	ccacctcgcc	6360
gcgcaggggc	tcgttgggtcc	agcagaggcg	cccgcccttg	cgcgagcaga	aggggggag	6420
cggttccagc	atgagctcgt	cggggggggtc	ggcgtccacg	gtgaagatgc	cgggcaggag	6480
ctcgggggtcg	aagtagctga	tgcagggtgcc	cagatcggtcc	agcgccgctt	gccagtcgcy	6540
cacggccagc	gcgcgctcgt	aggggctgag	gggcgtgccc	cagggcatgg	ggcgctgag	6600
cgcgaggcg	tacatgccgc	agatgtcgta	gacgtagagg	ggctcctcga	ggacgccgat	6660
gtagggtggg	tagcagcgcc	ccccgcggat	gctggcgcg	acgtagtcgt	acagctcgtg	6720
cgagggcgcg	aggagcccgg	tgccgagggt	ggagcgctgc	ggcttttcgg	cgcggtagac	6780
gatctggcg	aagatggcgt	gggagttgga	ggagatgggtg	ggcctctgga	agatgttgaa	6840
gtgggctg	ggcagtcga	ccgagtcctt	gatgaagtg	gcgtaggagt	cctgcagctt	6900
ggcgacgagc	tcggcggtga	cgaggacgtc	cagggcgag	tagtcgagg	tctcttgat	6960
gatgtcgta	ttgagctggc	ccttctgctt	ccacagctcg	cggttgagaa	ggaactcttc	7020
gcggtccttc	cagtactctt	cgagggggaa	cccgtcctga	tcggcacggt	aagagccac	7080
catgtagaac	tggttgacgg	ccttgtaggc	gcagcagccc	ttctccacgg	ggagggcgta	7140
agcttgcgcg	gccttgcgca	gggaggtgtg	ggtagggcg	aaggtgtcgc	gcaccatgac	7200
cttgaggaac	tggtgcttga	agtgcagggtc	gtcgcagccg	ccctgctccc	agagctggaa	7260
gtcgtgcgc	ttcttgtagg	cggggttggg	caaagcgaaa	gtaacatcgt	tgaagaggat	7320
cttgcccgcg	cggggcatga	agttgagagt	gatgcggaaa	ggctggggca	cctcgggccc	7380
gttggtgatg	acctgggagg	cgaggatgat	ctcgtcgaag	ccgttgatgt	tgtgcccagc	7440
gatgtagagt	tccacgaatc	gcgggcggcc	cttgacgtgg	ggcagcttct	tgagctcgtc	7500
gtagggtgagc	tcggcggggt	cgctgaggcc	gtgctgctcg	agggcccagt	cggcgagggtg	7560
ggggttggcg	ccgaggaagg	aagtccagag	atccacggcc	agggcggtct	gcaagcggtc	7620
ccggtactga	cggaactgct	ggcccaaggc	cattttttcg	ggggtgacgc	agtagaaggt	7680
gcgggggtcg	ccgtgccagc	ggccccactt	gagctggagg	gcgaggctgt	gggcgagctc	7740
gacgagcggc	gggtccccgg	agagtttcat	gaccagcatg	aaggggacga	gctgcttgcc	7800
gaaggacccc	atccaggtgt	aggtttccac	gtcgtagggtg	aggaagagcc	tttcgggtgcg	7860
aggatgcgag	ccgatgggga	agaactggat	ctcctgccac	cagttggagg	aatggctgtt	7920
gatgtgatgg	aagtagaaat	gccgacggcg	cgccgagcac	tcgtgcttgt	gtttatacaa	7980

gcgtccgcag tgctcgcaac gctgcacggg atgcacgtgc tgcacgagct gtacctgggt	8040
tcctttgacg aggaatttca gtgggcagtg gagcgctggc ggctgcatct ggtgctgtac	8100
tacgtcctgg ccatcggggt ggccatcgtc tgcctcgatg gtggctcatgc tgacgaggcc	8160
gcgcgggagg caggtccaga cctcggctcg gacgggtcgg agagcgagga cgagggcgcg	8220
caggccggag ctgtccaggg tcctgagacg ctgcggagtc aggtcagtgg gcagcggcgg	8280
cgcgcgggtg acttgcagga gcttttccag ggcgcgcggg aggtccagat ggtacttgat	8340
ctccacggcg ccgttggtgg cgacgtccac ggcttgacgg gtcccgtgcc cctggggcgc	8400
caccaccgtg ccccgtttct tcttggggtgc tggcgcgggc ggctccatgc ttagaagcgg	8460
cggcgaggac gcgcgcgggg cggcaggggc ggctcggggc ccggaggcag gggcggcagg	8520
ggcacgtcgg cgcgcgcgcg gggcagggtc tgggtactgcg ccggagagaag actggcgtga	8580
gcgacgacgc gacggttgac gtccctggatc tgacgcctct gggtgaaaggc cacgggaccc	8640
gtgagtttga acctgaaaga gagttcgaca gaatcaatct cggtatcggt gacggcgggc	8700
tgccgcagga tctcttgac gtgcgccgag ttgtcctggg aggcgatctc ggtcatgaac	8760
tgctcgatct cctcctcctg aaggctctcg cgaccggcgc gctcgacggg ggcgcgagg	8820
tcgttgagga tgcgggccat gagctgcgag aaggcggtca tgccggcctc gttccagacg	8880
cggctgtaga ccacggctcc gtcggggctcg cgcgcgcgca tgaccacctg ggcgagggtg	8940
agctcgacgt ggcgcgtgaa gaccgcgtag ttgcagaggc gctggtagag gtagttgagc	9000
gtggtggcga tgtgctcggg gacgaagaag tacatgatcc agcggcggag cggcatctcg	9060
ctgacgtcgc ccagggttc caagcgctcc atggcctcgt agaagtccac ggcgaagttg	9120
aaaaactggg agttgcgcgc cgagacggtc aactcctcct ccagaagacg gatgagctcg	9180
gcgatggtgg cgcgcacctc gcgctcgaag gccccggggg gctcctcttc ttccatctcc	9240
tcctcctctt ccatctcctc cactaacatc tcttctactt cctcctcagg aggcggcggc	9300
gggggagggg ccctgcgtcg ccggcgcgcg acgggcagac ggtcgatgaa gcgctcgatg	9360
gtctccccgc gccggcgacg catggtctcg gtgacggcgc gcccgctcctc ggggggcgcg	9420
agcgtgaaga cgcgcgcgcg catctccagg tggccgcggg ggggggtctcc gttgggcagg	9480
gagagggcgc tgacgatgca tcttatcaat tggcccgtag ggactccgcg caaggacctg	9540
agcgtctcga gatccacggg atccgaaaac cgctgaacga aggccttcgag ccagtcgcag	9600
tcgcaaggta ggctgagccc ggtttcttgt tcttcgggta tttggtcggg aggcgggcgg	9660

gcgatgctgc	tggtgatgaa	gttgaagtag	gcggtcctga	gacggcggat	ggtggcgagg	9720
agcaccaggt	ccttgggccc	ggcttgctgg	atgcgcagac	ggtcggccat	gccccaggcg	9780
tggtcctgac	acctggcgag	gtccttgtag	tagtcctgca	tgagccgctc	cacgggcacc	9840
tcctcctcgc	ccgcgcggcc	gtgcatgcgc	gtgagcccga	acccgcgctg	cggctggacg	9900
agcgccaggt	cggcgacgac	gcgtcggcg	aggatggcct	gctggatctg	ggtgaggggtg	9960
gtctggaagt	cgtcgaagtc	gacgaagcgg	tggtaggctc	cgggtgtgat	ggtgtaggag	10020
cagttggcca	tgacggacca	gttgacggtc	tggtggccgg	ggcgacagag	ctcgtggtac	10080
ttgaggcgcg	agtagggcg	cgtgtcgaag	atgtagtcgt	tgagggtgcg	cacgaggtac	10140
tggtatccga	cgaggaagtg	cggcggcgcc	tgccggtaga	gcggccatcg	ctcggtgccg	10200
ggggcgccgg	gcgcgaggtc	ctcgagcatg	aggcggtgg	agccgtagat	gtacctggac	10260
atccagggtga	tgccggcgcc	ggtggtggag	gcgcgcggga	actcgcgga	gcgggtccag	10320
atgttgcgca	gcggcaggaa	gtagttcatg	gtggccgcgg	tctggcccgt	gaggcgcgcg	10380
cagtcgtgga	tgctctagac	atacgggcaa	aaacgaaagc	ggtcagcgcc	tcgactccgt	10440
ggcctggagg	ctaagcgaac	gggttgggct	gcgcgtgtac	cccggttcga	gtccctgctc	10500
gaatcaggct	ggagccgcag	ctaactgggt	actggcactc	ccgtctcgac	ccaagcctgc	10560
taacgaaacc	tccaggatac	ggaggcggg	cgttttgccc	attttcgtca	ggccggaaat	10620
gaaactagta	agcgcgaaaa	gcggccgtcc	gcgatggctc	gctgccgtag	tctggagaaa	10680
gaatcgccag	ggttgcggtg	cgggtgtccc	cggttcgagc	ctcagcgctc	ggcgccggcc	10740
ggattccgcg	gctaactgtg	gcgtggctgc	cccgtcgttt	ccaagacccc	ttagccagcc	10800
gacttctcca	gttacggagc	gagccctct	ttttcttggt	tttttgccag	atgcattccg	10860
tactcgggca	gatgcgcccc	caccctccac	cacaaccgcc	cctaccgcag	cagcagcaac	10920
agccggcgct	tctgcccccg	ccccagcagc	agcagccagc	cactaccgcg	gcggccggcc	10980
tgagcggagc	cggcgttcag	tatgacctgg	ccttgggaaga	gggcgagggg	ctggcgcgcc	11040
tggggcgctc	gtcgccggag	cggcaccgcg	gcgtgcagat	gaaaagggag	gctcgcgagg	11100
cctacgtgcc	caagcagaac	ctgttcagag	acaggagcgg	cgaggagccc	gaggagatgc	11160
gcgcctcccc	cttccacgcg	gggcgggagc	tgcggcgcgg	cctggaccga	aagcgggtgc	11220
tgagggacga	ggatttcgag	gcggacgagc	tgacggggat	cagccccgcg	cgcgcgcacg	11280
tggccgcggc	caacctggtc	acggcgtagc	agcagaccgt	gaaggaggag	agcaacttcc	11340
aaaaatcott	caacaaccac	gtgcgcacgc	tgatcgcgcg	cgaggagggtg	accctggggc	11400

tgatgcacct	gtgggacctg	ctggaggcca	tcgtgcagaa	ccccacgagc	aagccgctga	11460
cggcgagct	gtttctggtg	gtgcagcaca	gtcgggacaa	cgagacgttc	agggaggcgc	11520
tgctgaatat	caccgagccc	gagggccgct	ggctcctgga	cctgggtaac	attctgcaga	11580
gcatcgtggt	gcaggagcgc	gggctgccgc	tgtccgagaa	gctggcggcc	atcaacttct	11640
cggtgctgag	cctgggcaag	tactacgcta	ggaagatcta	caagaccccc	tacgtgcccc	11700
tagacaagga	ggtgaagatc	gacgggtttt	acatgcgcat	gaccctgaaa	gtgctgaccc	11760
tgagcgacga	tctgggggtg	taccgcaacg	acaggatgca	ccgcgcgggtg	agcggcagcc	11820
gccggcgcg	gctgagcgac	caggagctga	tgcacagcct	gcagcggggc	ctgaccgggg	11880
ccgggaccga	gggggagagc	tactttgaca	tgggcgcgga	cctgcgctgg	cagcctagcc	11940
gccgggcctt	ggaagctgcc	ggcggttccc	cctacgtgga	ggaggtggac	gatgaggagg	12000
aggagggcga	gtacctggaa	gactgatggc	gcgaccgtat	ttttgctaga	tgcagcaaca	12060
gccaccgcg	cctcctgatc	ccgcgatgcy	ggcggcgctg	cagagccagc	cgtccggcat	12120
taactcctcg	gacgattgga	cccaggccat	gcaacgcata	atggcgctga	cgacccgcaa	12180
tcccgaagcc	tttagacagc	agcctcaggc	caaccgactc	tcggccatcc	tggaggccgt	12240
ggtgccctcg	cgctcgaacc	ccacgcacga	gaaggtgctg	gccatcgtga	acgcgctggt	12300
ggagaacaag	gccatccgcg	gcgacgaggc	cgggctgggt	tacaacgcgc	tgctggagcg	12360
cgtggcccg	tacaacagca	ccaacgtgca	gacgaacctg	gaccgcatgg	tgaccgacgt	12420
gcgcgaggcg	gtgtcgcagc	gcgagcgggt	ccaccgcgag	tcgaacctgg	gctccatggt	12480
ggcgctgaac	gccttcctga	gcacgcagcc	cgccaacgtg	ccccggggcc	aggaggacta	12540
caccaacttc	atcagcgcgc	tgcggctgat	ggtggccgag	gtgccccaga	gcgagggtgta	12600
ccagtcgggg	ccggactact	tcttcagac	cagtcgccag	ggcttgca	ccgtgaacct	12660
gagccaggct	ttcaagaact	tgacgggact	gtggggcggtg	caggccccgg	tcggggaccg	12720
cgcgacgggtg	tcgagcctgc	tgacgccgaa	ctcgcgcctg	ctgctgctgc	tggtggcgcc	12780
cttcacggac	agcggcagcg	tgagccgcga	ctcgtacctg	ggctacctgc	ttaacctgta	12840
ccgcgaggcc	atcgggcagg	cgacgtgga	cgagcagacc	taccaggaga	tcaccacagt	12900
gagccgcgcg	ctggggcagg	aggacccggg	caacctggag	gccaccctga	acttcctgct	12960
gaccaaccgg	tcgcagaaga	tcccgcacca	gtacgcgctg	agcaccgagg	aggagcgcac	13020
cctgcgctac	gtgcagcaga	gcgtggggct	gttcctgatg	caggaggggg	ccacgcccag	13080

cgccgcgctc gacatgaccg cgcgcaacat ggagcccagc atgtacgccc gcaaccgccc 13140
 gttcatcaat aagctgatgg actacttgca tcggggcgcc gccatgaact cggactactt 13200
 taccaacgcc atcttgaacc cgcactggct cccgccgccc gggttctaca cgggcgagta 13260
 cgacatgccc gacccaacg acgggttcct gtgggacgac gtggacagca gcgtgttctc 13320
 gccgcgcccc accaccacca ccgtgtggaa gaaagagggc ggggaccggc ggcgctcctc 13380
 ggcgctgtcc ggtcgcgcgg gtgctgccgc ggcggtgccc gaggccgcca gcccttccc 13440
 gagcctgccc ttttcgctga acagcgtgcg cagcagcgag ctgggtcggc tgaaggcgcc 13500
 gcgcctgctg ggcgaggagg agtacctgaa cgactccttg cttcggcccc agcgcgagaa 13560
 gaacttcccc aataacggga tagagagcct ggtggacaag atgagccgct ggaagacgta 13620
 cgcgcacgag cacagggacg agccccgagc tagcagcagc accggcgcca ccggtagacg 13680
 ccagcggcac gacaggcagc ggggtctggg gtgggacgat gaggattccg ccgacgacag 13740
 cagcgtgttg gacttgggtg ggagtgggtg tggtaaccgc ttcgctcacc tgcgcccccg 13800
 tatcgggcgc ctgatgtaag aatctgaaaa aataaaagac ggtactcacc aaggccatgg 13860
 cgaccagcgt gcgttcttct ctgttgtttg tagtagtatg atgaggcgcg tgtaccggga 13920
 gggctctcct ccctcgtacg agagcgtgat gcagcaggcg gtggcgggcg cgatgcagcc 13980
 cccgctggag gcgccttacg tgcccccgcg gtacctggcg cctacggagg ggcggaacag 14040
 cattcgttac tcggagctgg cacccttgta cgataccacc cggttgtacc tggtaggaaa 14100
 caagtgcgcg gacatcgctt cgtgaacta ccagaacgac cacagcaact tcctgaccac 14160
 cgtgggtgcag aacaacgatt tcacccccac ggaggccagc acccagacca tcaactttga 14220
 cgagcgctcg cgggtggggcg gccagctgaa aaccatcatg cacaccaaca tgcccaacgt 14280
 gaacgagttc atgtacagca acaagttcaa ggcgcgggtg atggctctgc gcaagacccc 14340
 caacggggtc acagtaacag atggtagtca ggacgagctg acctacgagt gggtaggagt 14400
 tgagctgccc gagggcaact tctcgtgac catgaccatc gatctgatga acaacgcat 14460
 catcgacaac tacttggcgg tggggcgcca gaacgggggtg ctggagagcg acatcgcgct 14520
 gaagtctgac acgcgcaact tccggctggg ctgggacccc gtgaccgagc tggtagtgcc 14580
 gggcgtgtac accaacgagg ccttcacccc cgacatcgtc ctgctgcccg gctgcggcgt 14640
 ggacttcacc gagagccgcc tcagcaacct gctgggcacg cgcaagcggc agcccttcca 14700
 ggagggcttc cagatcctgt acgaggacct ggaggggggc aacatccccg cgctgctgga 14760
 cgtggacgcc tacgagaaaa gcaaggagga tagcgccgcc gcggcgaccg cagccgtggc 14820

caccgcctct accgaggtgc ggggcgataa ttttgctagc gccgcgacac tggcagcggc 14880
cgaggcggct gaaaccgaaa gtaagatagt gatccagccg gtggagaagg acagcaagga 14940
gaggagctac aacgtgctcg cggacaagaa aaacaccgcc taccgcagct ggtaacctggc 15000
ctacaactac ggcgaccccg agaagggcgt gcgctcctgg acgtgctca ccacctcgga 15060
cgtcacctgc ggcgtggagc aagtctactg gtcgctgccc gacatgatgc aagaccgggt 15120
caccttcgcg tccacgcgtc aagttagcaa ctaccgggtg gtgggcgcgc agctcctgcc 15180
cgtctactcc aagagcttct tcaacgagca ggcgctctac tcgcagcagc tgcgcgcctt 15240
cacctcgctc acgcacgtct tcaaccgctt ccccgagaac cagatcctcg ttcgcccgc 15300
cgcgcccacc attaccaccg tcagtgaana cgctcctgct ctacagatc acgggaccct 15360
gccgctgcgc agcagtatcc ggggagtcca gcgcgtgacc gtcactgacg ccagacgccg 15420
cacctgcccc tacgtctaca aggccttggg cgtagtcgcg ccgcgcgtcc tctcgagccg 15480
caccttctaa aaaatgtcca ttctcatctc gccagtaat aacaccgggtt ggggcctgcg 15540
cgcgcccagc aagatgtacg gaggcgctcg ccaacgctcc acgcaacacc ccgtgcgcgt 15600
gcgcgggcac ttccgcgctc cctggggcgc cctcaagggc cgcgtgcgtc cgcgcaccac 15660
cgctgacgac gtgatcgacc aggtgggtggc gcagcgcgcg aactacacgc ccgcccgccg 15720
gcccgctctc accgtggacg ccgtcatcga cagcgtgggt gccgacgcgc gccggtacgc 15780
ccgcgccaaag agccggcggc ggcgcacgcg ccggcgccac cggagcacc ccgccatgcg 15840
cgcgggcgca gcccttgctg gcagggccag gcgcacggga cgcagggcca tgctcagggc 15900
ggccagacgc gcggcctccg gcagcagcag cgcgggcagg acccgagac gcgcggccac 15960
ggcgggcgcg gcggccatcg ccagcatgtc ccgcccgcgc cgcggcaacg tgtactgggt 16020
gcgcgacgcc gccaccgggt tgcgcgtgcc cgtgcgcacc cccccccctc gcaettgaag 16080
atgctgactt cgcgatgttg atgtgtccca gcggcgagga ggatgtccaa gcgcaaatcc 16140
aaggaaagaga tgctccaggt catcgcgcct gagatctacg gcccgggcgc ggtgaaggag 16200
gaaagaaagc cccgcaaact gaagcgggtc aaaaaggaca aaaaggagga ggaagatgtg 16260
gacggactgg tggagtgtgt gcgcgagttc gcccccggc ggcgcgtgca gtggcgcggg 16320
cggaaagtga aaccggtgct gcgaccgggc accacgggtg tcttcacgcc cggcgagcgt 16380
tccggctccg cctccaagcg ctctacgac gaggtgtacg gggacgagga catcctcgag 16440
caggcgcccg aacgtctggg cgagtttgct tacggcaagc gcagccgccc cgcgccttg 16500

aaagaggagg cggtgtccat cccgctggac cacggcaacc ccacgccgag cctgaagccg 16560
gtgacctgc agcaggtgct gcctgggtgcg gcgccgcgcc ggggcttcaa gcgcgagggc 16620
ggcgaggatc tgtacccgac catgcagctg atggtgccca agcgccagaa gctggaggac 16680
gtgctggagc acatgaaggt ggaccccgag gtgcagccc aggtcaaggt gcggcccatc 16740
aagcaggtgg ccccgggcct gggcgtgcag accgtggaca tcaagatccc cacggagccc 16800
atggaaacgc agaccgagcc cgtgaagccc agcaccagca ccatggagggt gcagacggat 16860
ccctggatgc cggcaccggc ttccaccacc cgccgaagac gcaagtacgg cggggccagc 16920
ctgctgatgc ccaactacgc gctgcatcct tccatcatcc ccacgccggg ctaccgcggc 16980
acgcgcttct accgcggcta caccagcagc cgccgccgca agaccaccac ccgccgccgc 17040
cgctgctcga cccgccgcag cagcaccgcg acttccgccg ccgccttggg gcggagagtg 17100
taccgcagcg ggcgcgagcc tctgacctg ccgcgcgcgc gctaccaccc gagcatcgcc 17160
atttaactac cgcctcctac ttgcagatat ggccctcaca tgccgcctcc gcgtcccat 17220
tacgggctac cgaggaagaa agccgcgccg tagaaggctg acggggaacg ggctgcgtcg 17280
ccatcaccac cggcgggcggc gcgccatcag caagcgggtg gggggaggct tcctgcccgc 17340
gctgatgccc atcatcgccg cggcgatcgg ggcgatcccc ggcatagctt ccgtggcggg 17400
gcaggcctct cagcgccact gagacacagc ttggaaaatt tgtaataaaa aatggactga 17460
cgctcctggg cctgtgatgt gtgttttttag atggaagaca tcaatttttc gtccctggca 17520
ccgcgacacg gcacgcggcc gtttatgggc acctggagcg acatcggcaa cagccaactg 17580
aacggggggcg ccttcaattg gagcagctctc tggagcgggc ttaagaattt cgggtccacg 17640
ctcaaaacct atggcaacaa ggcgtggaac agcagcacag ggcaggcgct gagggaag 17700
ctgaaagagc agaacttcca gcagaagggtg gtcgatggcc tggcctcggg catcaacggg 17760
gtggtggaac tggccaacca ggccgtgcag aaacagatca acagccgcct ggacgcggtc 17820
ccgcccgcg ggtccgtgga gatgcccag gtggaggagg agctgcctcc cctggacaag 17880
cgggcgaca agcgaccgcg tcccgacgcg gaggagacgc tgctgacgca cacggacgag 17940
ccgccccgt acgaggaggc ggtgaaactg ggtctgccc ccacgcggcc cgtgggcct 18000
ctggccaccg ggggtgctgaa acccagcagc agcagcagcc agcccgcgac cctggacttg 18060
cctccgctg ctccccgcc ctccacagtg gctaagcccc tgccgcgggt ggccgtcgcg 18120
tcgcgcgccc cccgaggccg ccccaggcg aactggcaga gcactctgaa cagcatcggt 18180
ggctctgggag tgcagagtgt gaagcgcgc cgctgctatt aaaagacact gtagcgctta 18240

acttgcttgt ctgtgtgtat atgtatgtcc gccgaccaga aggaggagga agaggcgct 18300
 cgccgagttg caagatggcc accdcatcga tgctgcccc gtgggcgtac atgcacatcg 18360
 ccggacagga cgcttcggag tacctgagtc cgggtctggt gcagttcgcc cgcgccacag 18420
 acacctactt cagtctgggg aacaagttta ggaacccac ggtggcgccc acgcacgatg 18480
 tgaccaccga ccgcagccag cggctgacgc tgcgcttcgt gcccggtggac cgcgaggaca 18540
 acacctactc gtacaaagtg cgctacacgc tggccgtggg cgacaaccgc gtgctggaca 18600
 tggccagcac ctactttgac atccgcggcg tgctggatcg gggccctagc ttcaaacctt 18660
 actccggcac cgcttacaac agcctggctc ccaagggagc gcccaacct tgccagtggg 18720
 catataaagc tgatgggtgat actggtacag aaaaaacct tacatatgga aatgcgcctg 18780
 tgcaaggcat tagtattaca aaagatggta ttcaacttgg aactgacact gatgatcagc 18840
 ccatttatgc agataaaact tatcaaccag agcctcaagt ggggtgatgct gaatggcatg 18900
 acatcactgg tactgatgaa aaatatggag gcagagctct caagcctgac accaaaatga 18960
 agccctgcta tggttctttt gccaaacct ccaataaaga aggaggtcag gcaaatgtga 19020
 aaaccgaaac aggcgggtacc aaagaatatg acattgacat ggcattcttc gataatcgaa 19080
 gtgcagctgc ggctggcctg gcccagaaa ttgttttgta tactgagaat gtggatctgg 19140
 aaactccaga tactcatatt gtatacaagg cgggcacaga tgacagcagc tcttctatca 19200
 atttgggtca gcagtcctat cccaacagac ccaactacat tggctttaga gacaacttta 19260
 tcgggctcat gtactacaac agcactggca acatgggcgt gctggctggg caggcctccc 19320
 agctgaatgc tgtgggtggac ttgcaggaca gaaacactga actgtcctac cagctcttgc 19380
 ttgactctct gggcgacaga accaggtatt tcagtatgtg gaatcaggcg gtggacagct 19440
 atgaccccca tgtgcgcatt attgaaaac acgggtgtgga ggatgaactc cctaactatt 19500
 gcttccccct ggatgctgtg ggtagaactg atacttacca gggaaattaag gccaatgggtg 19560
 ctgatcaaac cacctggacc aaagatgata ctgttaatga tgctaataaa ttgggcaagg 19620
 gcaatccttt cgccatggag atcaacatcc aggccaacct gtggcggaac ttcctctacg 19680
 cgaacgtggc gctgtacctg ccgactcct acaagtacac gccggccaac atcacgtctgc 19740
 cgaccaacac caacacctac gattacatga acggcgcgct ggtggcgccc tcgctggtgg 19800
 acgctacat caacatcggg gcgcgctggc cgctggaccc catggacaac gtcaaccctt 19860
 tcaaccacca ccgcaacgcg ggcctgcgct accgctccat gtcctgggc aacggcgct 19920

acgtgccctt ccacatccag gtgcccacaa agttcttgcg catcaagagc ctcttgcctc	19980
tgcccgggtc ctacacctac gagtgggaact tccgcaagga cgtcaacatg atcctgcaga	20040
gctccctcgg caacgacctg cgcacggacg gggcctccat cgccttcacc agcatcaacc	20100
tctacgccac cttcttcccc atggcgacac acaccgcctc cagctcagag gccatgctgc	20160
gcaacgacac caacgaccag tcttcaacg actacctctc ggcggccaac atgctctacc	20220
ccatcccggc caacgccacc aacgtgcca tctccatccc ctgcgcgaac tggggcgct	20280
tccgcgatg gtccttcacg cgcctcaaga cccgcgagac gccctcgctc ggctccgggt	20340
tcgacccta cttcgtctac tcgggctcca tcccctacct cgacggcacc ttctacctca	20400
accacacctt caagaaggct tccatcacct tcgactcctc cgtcagctgg cccggcaacg	20460
accgcctcct gacgccaac gagttcgaaa tcaagcgac cgtcgacgga gagggtaca	20520
acgtggccca gtgcaacatg accaaggact ggttcctggg ccagatgctg gccactaca	20580
acatcggcta ccagggttc tacgtgcccg agggctacaa ggaccgcatg tactccttct	20640
tccgcaactt ccagcccatg agcgcgcagg tcgtggacga ggtcaactac aaggactacc	20700
aggcgcctac cctggcctac cagcacaaca actcgggctt cgtcggctac ctgcgcca	20760
ccatgcgcca gggacagccc taccgcgcca actacccta cccgctcctc ggcaagagcg	20820
ccgtcgccag cgtcacccag aaaaagtctc tctgcgaccg ggtcatgtgg cgcctccct	20880
tctccagcaa cttcatgtcc atgggcgcgc tcaccgacct cggccagaac atgctctacg	20940
ccaactccgc ccacgcgcta gacatgaatt tcgaagtcga ccccatggat gagtccaccc	21000
ttctctatgt tgtcttcgaa gtcttcgacg tcgtccgagt gcaccagccc caccgcggcg	21060
tcacgagggc cgtctacctg cgcacgccct tctcggccgg caacgccacc acctaacccc	21120
cgtctttgct tcttgcaaga tgacggcctg tcggggctcc ggcgagcagg agctcagggc	21180
catctccgc gacctgggt gggggccctg ctctctgggc accttcgaca agcgttccc	21240
gggattcatg gcccgcaca agctggcctg cgccatcgtc aacacggccg gccgcgagac	21300
cgggggcgag cactggetgg ccttcgcctg gaaccgcgc tcccacacct gctacctctt	21360
cgacccttc gggttctcgg acgagcgct caagcagatc taccagtctg agtacgagg	21420
cctgctgcgc cgcagcgccc tggccaccga ggaccgctgc gtcacctgg aaaagtccac	21480
ccagaccgtg cagggtccgc gtcggccgc ctgcgggctc ttctgctgca tgttctgca	21540
cgccttcgtg cactggcccg accgccccat ggacaagaac cccaccatga acttgctgac	21600
gggggtgccc aacggcatgc tccagtcgcc ccaggtggaa cccacctgc gccgcaacca	21660

ggaggcgctc	taccgcttcc	tcaacgcccc	ctcgcctac	tttcgctccc	accgcgcgcg	21720
catcgagaag	gccaccgcct	tcgaccgcat	gaatcaagac	atgtaaaccg	tgtgtgtatg	21780
tgaatgcttt	attcataata	aacagcacat	gtttatgcc	ccttttctga	ggctctgact	21840
ttatttagaa	atcgaagggg	ttctgccggc	tctcggcggtg	ccccgcgggg	agggatacgt	21900
tgcggaactg	gtacttgggc	agccacttga	actcggggat	cagcagcttc	ggcacgggga	21960
ggtcggggaa	cgagtcgctc	cacagcttgc	gcgtgagttg	cagggcgccc	agcaggtcgg	22020
gcgcggagat	cttgaaatcg	cagttgggac	ccgcgttctg	gcgcggggag	ttgcggtaca	22080
cggggttgca	gcactggaac	accatcaggg	ccgggtgctt	cacgctcgcc	agcaccgtcg	22140
cgtcggtgat	gccctccacg	tccagatcct	cggcgttggc	catcccgaag	ggggtcatct	22200
tgcaggctctg	ccgccccatg	ctgggcacgc	agccgggctt	gtggttgcaa	tcgcagtgc	22260
gggggatcag	catcatctgg	gcctgctcgg	agctcatgcc	cgggtacatg	gccttcatga	22320
aagcctccag	ctggcggaag	gcctgctcgg	ccttgccgcc	ctcgggtgaag	aagacccccg	22380
aggacttgct	agagaactgg	ttgggtggcg	agccggcgct	gtgcacgcag	cagcgcgct	22440
cgttggtggc	cagctgcacc	acgctgcgcc	cccagcggtt	ctgggtgatc	ttggcccggg	22500
cggggttctc	cttcagcgcg	cgctgcccg	tctcgcctgc	cacatccatc	tcgatcgtgt	22560
gctccttctg	gatcatcacg	gtcccgtgca	ggcatcgag	cttgccctcg	gcctcgggtgc	22620
accggtgcag	ccacagcgcg	cagccgggtgc	actcccagtt	cttggtggcg	atctgggagt	22680
gcgagtgcac	gaagccctgc	aggaagcggc	ccatcatcgt	ggtcagggtc	ttgttgetgg	22740
tgaaggtcag	cgggatgccg	cggtgctcct	cgttcacata	cagggtggcag	atgcggcggt	22800
acacctcgcc	ctgctcgggc	atcagctgga	aggcggactt	caggctcgctc	tccatgcgggt	22860
accggtccat	cagcagcgct	atgacttcca	tgcccttctc	ccaggccgag	acgatcggca	22920
ggctcagggg	gttcttcacc	gccgttgtea	tcttagtcgc	cgccgctgag	gtcagggggg	22980
cgttctcgct	caggggtctca	aacactcgct	tgccgtcctt	ctcgggtgatg	cgcaacggggg	23040
gaaagctgaa	gcccacggcc	gccagctcct	cctcggcgctg	cctttctgctc	tcgctgtcct	23100
ggctgatgtc	ttgcaaaggc	acatgcttgg	tcttgcgggg	tttctttttg	ggcggcagag	23160
gcggcgggcg	agacgtgctg	ggcgagcgcg	agttctcgct	caccacgact	atttcttctt	23220
cttgcccgct	gtccgagacc	acggggcggt	aggcatgcct	cttctggggc	agaggcgag	23280
gcgacgggct	ctcgcggttc	ggcggggcg	tggcagagcc	ccttccgcgt	tcgggggtgc	23340

gctcctggcg gcgctgctct gactgacttc ctccgcggcc ggccattgtg ttctcctagg 23400
gagcaacaag catggagact cagccatcgt cgccaacatc gccatctgcc cccgcgcgcg 23460
ccgacgagaa ccagcagcag aatgaaagct taaccgcccc gccgcccagc cccacctccg 23520
acgccgcgcg gggcccagac atgcaagaga tggaggaatc catcgagatt gacctgggct 23580
acgtgacgcc cgcggagcac gaggaggagc tggcagcgcg cttttcagcc ccggaagaga 23640
accaccaaga gcagccagag caggaagcag agagcgagca gcagcaggct gggctcgagc 23700
atggcgacta cctgagcggg gcagaggacg tgctcatcaa gcctctggcc cgccaatgca 23760
tcctcgtcaa ggacgcgctg ctccgaccgc ccgaggtgcc cctcagcgctg gcggagctca 23820
gccgcgccta cgagcgcaac ctcttctcgc cgcgcgtgcc ccccaagcgc cagcccaacg 23880
gcacctgcca gccaacccg cgcctcaact tctaccgggt ctccgcgggtg cccgaggccc 23940
tgccaccta ccacctcttt ttcaagaacc aaaggatccc cgtctcctgc cgcgccaacc 24000
gcaccgcgcg cgacgccctg ctcaacctgg gtcccggcgc ccgcctacct gatatcgct 24060
ccttggaaga ggttcccaag atcttcgagg gtctgggcag cgacgagact cgggcccgcga 24120
acgctctgca aggaagcggg gaggagcatg agcaccacag cgccctgggtg gagttggaag 24180
gcgacaacgc gcgcctggcg gtgctcaagc gcacggtcga gctgaccac ttcgcctacc 24240
cggcgctcaa cctgcccccc aaggctcatga gcgcgctcat ggaccagggtg ctcatcaagc 24300
gcgcctcgcc cctctcggtt gaggacatgc aggaccccga gagctcggac gagggcaagc 24360
ccgtggtcag cgacgagcag ctggcgcgct ggctgggagc gagtagcacc cccagagct 24420
tggaagagcg gcgcaagctc atgatggccg tggctcctggt gaccgtggag ctggagtgct 24480
tgcccgctt cttcgccgac gcagagaccc tgcgcaagggt cgaggagaac ctgcactacc 24540
tcttcaggca cgggtttgtg cgccaggcct gcaagatctc caacgtggag ctgaccaacc 24600
tggtctccta catgggcatc ctgcacgaga accgcctggg gcagaacgtg ctgcacacca 24660
ccctgcgcgg ggaggcccgc cgcgactaca tccgcgactg cgtctacctg tactctgcc 24720
acacctggca gacgggcatg ggcggtgtggc agcagtgcct ggaggagcag aacctgaaag 24780
agctctgcaa gctcctgcag aagaacctga aggccctgtg gaccgggttc gacgagcgca 24840
ccaccgcctc ggacctggcc gacctcatct tccccgagcg cctgcggctg acgctgcgca 24900
acggactgcc cgactttatg agtcaaagca tgttgcaaaa ctttcgctct tcatcctcg 24960
aacgctccgg gatcctgccc gccacctgct ccgcgctgcc ctccgacttc gtgccgctga 25020
ccttcgcga gtgccccccg ccgctctgga gccactgcta cctgctgcgc ctggccaact 25080

acctggccta ccactcggac gtgatcgagg acgtcagcgg cgagggtctg ctcgagtgcc 25140
actgccgctg caacctctgc acgccgcacc gctccctggc ctgcaacccc cagctgctga 25200
gcgagaccca gatcatcggc adcttcgagt tgcaaggccc cggcgagggc aaggggggtc 25260
tgaaactcac cccggggctg tggacctcgg cctacttgcg caagtctgtg cccgaggact 25320
accatccctt cgagatcagg ttctacgagg accaatccca gccgcccaag gccgaactgt 25380
cggcctgcgt catcaccag ggggccatcc tggcccaatt gcaagccatc cagaaatccc 25440
gccaagaatt tctgctgaaa aagggccacg ggggtctacct ggacccccag accggagagg 25500
agctcaaccc cagcttcccc caggatgccc cgaggaagca gcaagaagct gaaagtggag 25560
ctgccgccgc cggaggattt ggaggaaagac tgggagagca gtcaggcaga ggaggaggag 25620
atggaagact gggacagcac tcaggcagag gaggacagcc tgcaagacag tctggaagac 25680
gaggtggagg agggaggcaga ggaagaagca gccgccgcca gaccgtcgtc ctccggcgag 25740
aaagcaagca gcacggatac catctccgct ccgggtcggg gtcgcggcga ccggggccac 25800
agtaggtggg acgagaccgg gcgcttcccg aaccccacca ccagaccgg taagaaggag 25860
cggcagggat acaagtcctg gcgggggcac aaaaacgcca tcgtctcctg cttgcaagcc 25920
tgcgggggca acatctcctt caccggccgc tacctgctct tccaccgagg ggtgaacttc 25980
ccccgcaaca tcttgcatca ctaccgtcac ctccacagcc cctactactg tttccaagaa 26040
gaggcagaaa cccagcagca gcagaaaacc agcggcagca gcagctagaa aatccacagc 26100
ggcgccaggt ggactgagga tcgcagcgaa cgagccggcg cagaccgggg agctgaggaa 26160
ccggatcttt cccacctctt atgcatctt ccagcagagt cgggggcagg agcaggaaact 26220
gaaagtcaag aaccgttctc tgcgtctgct caccgcagt tgtctgtatc acaagagcga 26280
agaccaactt cagcgcactc tcgaggacgc cgaggctctc ttcaacaagt actgcgcgct 26340
cactcttaaa gagtagcccg cgcccgccca cacacggaaa aaggcgggaa ttacgtcacc 26400
acctgcgccc ttgcgccgac catcatcatg agcaaagaga ttcccacgcc ttacatgtgg 26460
agctaccagc cccagatggg cctggccgcc ggcgccgcc aggactactc caccgcgatg 26520
aactgggtca gcgcggggcc cgcgatgac tcacgggtga atgacatccg cgcccgccga 26580
aaccagatac tcctagaaca gtcagcgatc accgccacgc cccgccatca ccttaatccg 26640
cgtaattggc ccgcccctt ggtgtaccag gaaattcccc agcccacgac cgtactactt 26700
ccgcgagacg cccaggccga agtccagctg actaactcag gtgtccagct ggccggcgcc 26760

gocgccctgt	gtcgtcaccg	ccccgctcag	ggatataaagc	ggctgggtgat	ccgagggcaga	26820
ggcacacagc	tcaacgacga	gggtggtagc	tcttcgctgg	gtctgcgacc	tgacggagtc	26880
ttccaactcg	ccggatcggg	gagatcttcc	ttcacgcctc	gtcaggccgt	cctgactttg	26940
gagagttcgt	cctcgcagcc	ccgctcgggt	ggcatcggca	ctctccagtt	cgtggaggag	27000
ttcactccct	cggtctactt	caacccttc	tccggctccc	ccggccacta	cccggacgag	27060
ttcatcccga	acttcgacgc	catcagcgag	tgggtggacg	gctacgattg	aatgtcccat	27120
gggtggcgag	ctgacctagc	tgggttcga	cacctggacc	actgccgcgc	cttcgcgtgc	27180
ttcgctcggg	atctcgccga	gtttgcctac	tttgagctgc	ccgaggagca	ccctcagggc	27240
ccggcccacg	gagtgcggat	catcgtcgaa	gggggcctcg	actccacct	gcttcggatc	27300
ttcagccagc	gaccgatcct	ggtcgagcgc	gagcaaggac	agacccttct	gaccctgtac	27360
tgcattctga	accaccccg	cctgcataaa	agtccttctg	gtctgctgtg	tactgagtat	27420
aataaaaagct	gagatcagcg	actactccgg	actcgattgt	gggtgttcctg	ctatcaaccg	27480
gtccctgttc	ttcacccgga	acgagaccga	gtccagctt	cagtgtgaagc	cccacaagaa	27540
gtacctcacc	tggctgttcc	agggtcctcc	gatcgccgtt	gtcaaccact	gcgacaacga	27600
cggagtcctg	ctgagcggcc	ccgccaacct	tactttttcc	acccgcagaa	gcaagctcca	27660
gctcttccaa	cccttcctcc	ccgggacctc	tcagtgcgtc	tcgggaccct	gccatcacac	27720
cttcacactg	atcccgata	ccacagcgcc	gtcccccgt	actaacaacc	aaactaccca	27780
ccatcgccac	cgtcgcgacc	tttctgaatc	taacactacc	acccacaccg	gaggtgagct	27840
ccgaggtcga	ccaacctctg	ggatttacta	cggccctcgg	gaggtgggtg	ggttaatagc	27900
gctaggccta	gttggtgggtg	ggcttttggc	tctctgtctac	ctatacctcc	cttgctgttc	27960
gtacttagtg	gtgctgtgtt	gctgggttaa	gaaatgggga	agatcaccct	agtgagctgc	28020
gggtgcgctgg	tggcggtggt	gggtgttttcg	attgtgggac	tgggcggcgc	ggctgtagtg	28080
aaggagaagg	ccgatccctg	cttgcatctc	aatcccgaca	attgccagct	gagttttcag	28140
cccgatggca	atcgggtgcgc	gggtgctgac	aagtgcggat	gggaatgcga	gaacgtgaga	28200
atcgagtaca	ataacaagac	tcggaacaat	actctcgcgt	ccgtgtggca	gcccggggac	28260
cccgagtggg	acaccgtctc	tgtccccggg	gtgacgggt	ccccgcgcac	cgtgaacaat	28320
acttttcattt	ttgcgcacat	gtgcgacacg	gtcatgtgga	tgagcaagca	gtacgatatg	28380
tggcccccca	cgaaggagaa	catcgtgggc	ttctccatcg	cttacagcgc	gtgcacggcg	28440
ctaatacccg	ctatcgtgtg	cctgagcatt	cacatgctca	tcgctattcg	cccagaaat	28500

aatgccgaaa aagagaaaca gccataacac gttttttcac acaccttttt cagaccatgg 28560
cctctgttaa attttttgctt ttatttgcca gtctcattac tgttataagt aatgagaaac 28620
tcactattta cattggcact aaccacactt tagacggaat tccaaaatcc tcatgggtatt 28680
gctattttga tcaagatcca gacttaacta tagaactgtg tggtacaag ggaaaaaata 28740
caagcattca tttaattaac tttaattgcg gagacaattt gaaattaatt aatatcacta 28800
aagagtatgg aggtatgtat tactatgttg cagaaaataa caacatgcag ttttatgaag 28860
ttactgtaac taatcccacc acacctagaa caacaacaac caccaccaca aaaactacac 28920
ctgttaccac tatgcagctc actaccaata acatttttgc catgcgtcaa atgggtcaaca 28980
atagcactca acccacccca cccagtggagg aaattcccaa atccatgatt ggcattattg 29040
ttgtctgagt ggtgtgcatg ttgatcatcg ccttgtgcat ggtgtactat gccttctgct 29100
acagaaagca cagactgaac gacaagctgg aacacttact aagtgttgaa ttttaatttt 29160
ttagaaccat gaagatccta ggccttttaa tttttctat cattacctct gctctatgca 29220
attctgacaa tgaggacgtt actgtcggtg tcggaaccaa ttatacactg aaaggtccag 29280
cgaagggtat gctttcgtgg tattgctggg ttggaactga cgagcaacag acagagctct 29340
gcaatgctca aaaaggcaaa acctcaaatt ctaaaatctc taattatcaa tgcaatggca 29400
ctgacttagt actgctcaat gtcacgaaag catatgctgg cagctacacc tgccctggag 29460
atgatactga gaacatgatt ttttacaag tggaaagtggt tgatccact actccacctc 29520
caccaccac aactactcac accacacaca cagaacaac cacagcagag gaggcagcaa 29580
agttagcctt gcaggccaa gacagttcat ttgttggcat taccctaca cctgatcagc 29640
ggtgtccggg gctgctcgtc agcggcattg tcggtgtgct ttcgggatta gcagtcataa 29700
tcactgcat gttcattttt gcttgtgct atagaaggct ttaccgacaa aatcagacc 29760
cactgctgaa cctctatggt taattttttc cagagccatg aaggcagtta gcactctagt 29820
tttttgtct ttgattggca ctgttttttag tgtagcttt ttgaaacaaa tcaatgttac 29880
tgagggggaa aatgtgacac tggtaggcgt agagggtgct caaaatacca cctggacaaa 29940
attccatcta gatgggtgga aagaaatttg cacctggaat gtcagtactt atacatgtga 30000
aggagttaat cttaccattg tcaatgtcag ccaaattcaa aagggttga ttaaagggca 30060
atctgttagt gtagcaata gtgggtacta taccagcat actcttatct atgacattat 30120
agttatacca ctgcctacac ctagcccacc tagcactacc acacagacaa cccacactac 30180

acaaacaacc	acatacagta	catcaaata	gcctaccacc	actacaacag	cagaggttgc	30240
cagctcgtct	ggggtccgag	tggcattttt	gatgttggcc	ccatctagca	gtccactgc	30300
tagtaccaat	gagcagacta	ctgaattttt	gtccactgtc	gagagccaca	ccacagctac	30360
ctcgagtgc	ttctctagca	cggccaatct	atcctcgtct	tcctctacac	caatcagtcc	30420
cgctactact	cctacccccg	ctattctccc	cactccccctg	aagcaaacag	acggcgacat	30480
gcaatggcag	atcacccctgc	tcattgtgat	cgggttggtc	atcctggccg	tgttgcctta	30540
ctacatcttc	tgcgcgcgca	ttcccaacgc	gcaccgcaag	cgggcctaca	agcccatcgt	30600
tgtcgggag	cgggagccgc	ttcaggtgga	aggggtgcta	aggaatcttc	tcttctcttt	30660
tacagtatgg	tgattgaatt	atgattccta	gacaaatctt	gatcactatt	cttatctgcc	30720
tcctccaagt	ctgtgccacc	ctcgtctctg	tggccaacgc	cagtccagac	tgtattgggc	30780
ccttcgcctc	ctacgtgctc	tttgccctca	tcacctgcat	ctgctgctgt	agcctagtct	30840
gcctgcttat	cacctctctc	cagttcattg	actggatctt	tgtgcgcac	gcctacctgc	30900
gccaccaccc	ccagtaccgc	gaccagcgag	tggcgcggt	gctcaggatc	ctctgataag	30960
catgcggtct	ctgctacttc	tcgcgtctct	gctgttagtg	ctccccctc	ccgtcgaccc	31020
cgggaccccc	accagtcctc	cggaggaggt	cgcgaatgc	aaattccaag	aacctggaa	31080
attcctcaaa	tgctaccgcc	aaaaatcaga	catgcatccc	agctggatca	tgatcattgg	31140
gatcgtgaac	attctggcct	gcacctcat	ctcctttgtg	atttacctct	gctttgactt	31200
tggttggaac	tcgccagagg	cgctctatct	cccgcctgaa	cctgacacac	caccacagca	31260
acctcaggca	cacgcactac	caccaccacc	acagcctagg	ccacaataca	tgcccatatt	31320
agactatgag	gccgagccac	agcgacccat	gctccccgct	attagttact	tcaatctaac	31380
cggcgagagat	gactgaccca	ctggccaaca	acaacgtcaa	cgaccttctc	ctggacatgg	31440
acggccgcgc	ctcggagcag	cgactcgccc	aacttcgcat	tcgccagcag	caggagagag	31500
ccgtcaagga	gctgcaggac	ggcatagcca	tccaccagtg	caagaaaggc	atcttctgcc	31560
tggtgaaaca	ggccaagatc	tcctacgagg	tcaccacagc	cgaccatcgc	ctctcctacg	31620
agctcctgca	gcagcgccag	aagttcacct	gcctgggtcg	agtcaacccc	atcgtcatca	31680
cccagcagtc	gggcgatacc	aaggggtgca	tccactgtct	ctgcgactcc	cccgactgcg	31740
tccacactct	gatcaagacc	ctctgcggcc	tccgcgacct	cctccccatg	aactaatcac	31800
ccccttatcc	agtgaataaa	agatcatatt	gatgatttga	gtttaataaa	aataaagaat	31860
cacttacttg	aatctgata	ccaggtctct	gtccatgttt	tctgccaaca	ccaattcact	31920

ccccctcttcc	cagctctggt	actgcaggcc	ccggcgggct	gcaaacttcc	tccacaccct	31980
gaaggggatg	tcaaattcct	cctgtccctc	aatcttcatt	ttatcttcta	tcagatgtcc	32040
aaaaagcgcg	tccgggtgga	tgatgacttc	gaccccgctc	accctacga	tgagacaac	32100
gcaccgaccg	tgcccttcat	caaccccccc	ttcgtctctt	cagatggatt	ccaagagaag	32160
cccctggggg	tgctgtccct	gcgtctggcc	gatcccgta	ccaccaagaa	cggggaaatc	32220
accctcaagc	tgggagatgg	ggtggacctc	gactcctcgg	gaaaactcat	ctccaacacg	32280
gccaccaagg	ccgccgcccc	tctcagtttt	tccaacaaca	ccatttcctt	taacatggat	32340
accctttttt	acaacaacaa	tggaaagtta	ggcatgaaag	tcactgtctc	actgaagata	32400
ctagacacag	acttgctaaa	aacacttggt	gtagcttatg	gacaagggtt	aggaacaaac	32460
accactggtg	cccttggttc	ccaactagca	tccccacttg	cttttgatag	caatagcaaa	32520
attgccctta	atttaggcaa	tggaccattg	aaagtggatg	caaataagact	gaacatcaat	32580
tgcaatagag	gactctatgt	tactaccaca	aaagatgcac	tggaagccaa	tataagttag	32640
gctaattgta	tgacatttat	aggaaatgcc	atgggtgtca	atattgatac	acaaaaaggc	32700
ttgcaatttg	gcaccactag	taccgtcgca	gatgttaaaa	acgcttacct	catacaaatc	32760
aaacttgtag	ctggtctcac	atltgacagc	acaggtgcaa	ttggtgcatg	gaacaaagat	32820
gatgacaagc	ttacactatg	gaccacagcc	gacccctctc	caaattgtca	catatattct	32880
gaaaaggatg	ctaagcttac	actttgcttg	acaaagtgtg	gcagtcagat	tctgggcact	32940
gtttccctca	tagctgttga	tactggcagt	ttaaatccca	taacaggaac	agtaaccact	33000
gctcttgtct	cacttaaatt	cgatgcaaat	ggagttttgc	aaagcagctc	aacactagac	33060
tcagactatt	ggaatttcag	acaggggagat	gttacacctg	ctgaagccta	tactaatgct	33120
atagggtttca	tgcccaatct	aaaagcatac	cctaaaaaca	caagtggagc	tgcaaaaagt	33180
cacattgttg	ggaaagtgtg	cctacatggg	gatacaggca	aaccactgga	cctcattatt	33240
actttcaatg	aaacaagtga	tgaatcttgc	acttactgta	ttaactttca	atggcagtgg	33300
ggggctgatc	aataataaaa	tgaaacactt	gccgtcagtt	cattcacctt	ttcctatatt	33360
gctaaagaat	aaacccact	ctgtacccca	tctctgtcta	tggaaaaaac	tctgaaacac	33420
aaaataaaat	aaagttcaag	tggttttattg	attcaacagt	tttacaggat	tcgagcagtt	33480
atttttcctc	caccctccca	ggacatggaa	tacaccacce	tctccccccg	cacagccttg	33540
aacatctgaa	tgccattggt	gatggacatg	cttttggtct	ccacgttcca	cacagtttca	33600

gagcgagcca	gtctcgggtc	ggtcágggag	atgaaaccct	ccgggcactc	ccgcattctgc	33660
acctcacagc	tcaacagctg	aggattgtcc	tcggtggtcg	ggatcacggg	tatctggaag	33720
aagcagaaga	gcggcgggtg	gaatcatagt	ccgcgaacgg	gatcgggccg	tgggtgctgca	33780
tcaggccccg	cagcagtcgc	tgctgcgcgc	gctccgtcaa	gctgctgctc	aggggggtccg	33840
ggtcacagga	ctccctcagc	atgatgcccc	cggccctcag	catcagtcgt	ctgggtgcggc	33900
gggcgcagca	gcgcatgcgg	atctcgctca	ggtcgctgca	gtacgtgcaa	cacaggacca	33960
ccaggttgtt	caacagtcca	tagttcaaca	cgtccagcc	gaaactcctc	gcgggaagga	34020
tgctaccac	gtggccgtcg	taccagatcc	tcaggtaa	caagtggcgc	ccctccaga	34080
acacgctgcc	catgtacatg	atctccttgg	gcattgtggc	gttcaccacc	tcccggtacc	34140
acatcacctt	ctgggtgaac	atgcagcccc	ggatgatcct	gcggaaccac	agggccagca	34200
ccgccccgcc	cgccatgcag	cgaagagacc	ccgggtcccg	acaatggcaa	tggaggaccc	34260
accgctcgta	cccgtggatc	atctgggagc	tgaacaagtc	tatgttggca	cagcacaggc	34320
atatgctcat	gcattctctt	agcactctca	gctcctcggg	ggtcaaaacc	atatccagg	34380
gcacggggaa	ctcttgacag	acagcgaacc	ccgcagaaca	gggcaatcct	cgcatataac	34440
ttacattgtg	catggacagg	gtatcgcaat	caggcagcac	cgggtgatcc	tccaccagag	34500
aagcgcgggt	ctcgggtctc	tcacagcgtg	gtaagggggc	cggccgatac	gggtgatggc	34560
gggacgcggc	tgatcggtgt	cgcgaccgtg	ttatgatgca	gttgctttcg	gacattttcg	34620
tacttgctgt	agcagaacct	ggtcggggcg	ctgcacaccg	atcgccggcg	gcgggtcccg	34680
cgttggaac	gctcgggtgt	gaagttgtaa	aacagccact	ctctcagacc	gtgcagcaga	34740
tctagggcct	caggagtgat	gaagatcccc	tcattgcctga	tggctcta	catatcgacc	34800
accgtggaat	gggacagacc	cagccagatg	atgcaatttt	gttgggtttc	ggtgacggcg	34860
ggggagggaa	gaacaggaag	aaccatgatt	aactttta	ccaaacggtc	tcggagcact	34920
tcaaatgaa	gatcgcgagg	atggcacctc	tcgccccgc	tgtgttggtg	gaaaataaca	34980
gccagggtcaa	aggtgatacg	gttctcgaga	tgttccacgg	tggcttcag	caaagcctcc	35040
acgcgcacat	ccagaaacaa	gacaatagcg	aaagcgggag	ggttctctaa	ttcctcaatc	35100
atcatgttac	actcctgcac	catccccaga	taattttcat	ttttccagcc	ttgaatgatt	35160
cgaactagtt	cctgaggtaa	atccaagcca	gccatgataa	agagctcgcg	cagagcgccc	35220
tccaccggca	ttcttaagca	cacctcata	attccaagat	attctgctcc	tggttcacct	35280
gcagcagatt	gacaagcgga	atatcaaat	ctctgcgcgc	atccctaagc	tcctccctca	35340

gcaataactg taagtactct ttcatactct ctccgaaatt tttagccata ggaccaccag 35400
gaataagatt agggcaagcc acagtacaga taaaccgaag tcctccccag tgagcattgc 35460
caaatgcaag actgctataa gcâtgctggc tagaccgggt gatatcttcc agataactgg 35520
acagaaaatc gcccaggcaa tttttaagaa aatcaacaaa agaaaaatcc tccagggtgca 35580
cgttttagagc ctcggaaca acgatggagt aaatgcaagc ggtgcgttcc agcatgggta 35640
gttagctgat ctgtagaaaa aaacaaaaat gaacattaaa ccatgctagc ctggcgaaca 35700
ggtagggtaaa tcgttctctc cagcaccagg caggccacgg ggtctccggc acgaccctcg 35760
taaaaattgt cgctatgatt gaaaaccatc acagagagac gttcccgggtg gccggcgtga 35820
atgattcgac aagatgaata cacccccga acattggcgt ccgcgagtga aaaaaagcgc 35880
ccaaggaagc aataaggcac tacaatgctc agtctcaagt ccagcaaagc gatgccatgc 35940
ggatgaagca caaaattctc aggtgcgtac aaaatgtaat tactcccctc ctgcacaggc 36000
agcaaagccc ccgatccctc caggtagaca taaaagcct cagcgtccat agcttaccga 36060
gcagcagcac acaacaggcg caagagtcag agaaaggctg agctctaacc tgtccacccg 36120
ctctctgctc aatatatagc ccagatctac actgacgtaa aggccaaagt ctaaaaatac 36180
ccgccaaata atcacacacg ccagcacac gcccagaaac cggtagacaca ctcaaaaaaa 36240
tacgcgcact tcctcaaagc cccaaactgc cgtcatttcc ggggtccac gctacgtcat 36300
caaaattcga ctttcaaatt ccgtcgaccg ttaaaaacgt cggccgcccc gccctaacg 36360
gtcgcgctc ccgcagcaa tcaccgcccc gcatcccaa attcaaatac ctcatttgca 36420
tattaacgcg caccaaaagt ttgaggtata ttattgatga tg 36462

<210> 2
<211> 36604
<212> DNA
<213> chimpanzee adenovirus serotype Pan6

<400> 2
catcatcaat aatatacctc aaacttttgg tgcgcgttaa tatgcaaag agctgtttga 60
atttggggag ggaggaagg gattggctgc gggagcggcg accgttaggg gcggggcggg 120
tgacgttttg atgacgtggc tatgaggcgg agcgggtttg caagttctcg tgggaaaagt 180
gacgtcaaac gaggtgtggg ttgaacacgg aaataactcaa tttcccgcg ctctctgaca 240
ggaaatgagg tgtttctggg cggatgcaag tgaaaacggg ccattttcgc gcgaaaactg 300
aatgaggaag tgaaaatctg agtaatttcg cgtttatggc agggaggagt atttgccgag 360

ggccgagtag	actttgaccg	attacgtggg	ggtttcgatt	accgtatttt	tcacctaaat	420
ttccgcgtac	ggtgtcaaag	tccggtgttt	ttacgtagge	gtcagctgat	cgccagggta	480
tttaaacctg	cgctctctag	tcaagaggcc	actcttgagt	gccagcgagt	agagttttct	540
cctccgcgcc	gcgagtcaga	tctacacttt	gaaagatgag	gcacctgaga	gacctgcccc	600
gtaatgtttt	cctggctact	gggaacgaga	ttctggaatt	ggtgggtggac	gccatgatgg	660
gtgacgaccc	tccagagccc	cctaccccat	ttgaggcgcc	ttcgctgtac	gatttgatat	720
atctggaggt	ggatgtgccc	gagagcgacc	ctaacgagga	ggcgggtgaat	gatttgttta	780
gcgatgccgc	gctgctggct	gccgagcagg	ctaatacgga	ctctggctca	gacagcgatt	840
cctctctcca	taccccgaga	cccggcagag	gtgagaaaaa	gatccccgag	cttaaagggg	900
aagagctcga	cctgcgctgc	tatgaggaat	gcttgccctc	gagcgatgat	gaggaggacg	960
aggaggcgat	tcgagctgcg	gtgaaccagg	gagtgaaaac	tgcggggcgag	agcttttagcc	1020
tggactgtcc	tactctgccc	ggacacggct	gtaagtcttg	tgaatttcat	cgcatgaata	1080
ctggagataa	gaatgtgatg	tgtgccctgt	gctatatgag	agcttacaac	cattgtgttt	1140
acagtaagtg	tgattaactt	tagttgggaa	ggcagagggg	gactgggtgc	tgactggttt	1200
atztatgtat	atgttttttt	atgtgtaggt	cccgctctctg	acgtagatga	gacccccact	1260
tcagagtgca	tttcatcacc	ccagaaaatt	ggcgaggaac	cgcccgaaga	tattattcat	1320
agaccagttg	cagtgcgagt	caccgggcgg	agagcagctg	tggagagttt	ggatgacttg	1380
ctacaggggtg	gggatgaacc	tttggacttg	tgtaccggga	aacgccccag	gcactaagtg	1440
ccacacatgt	gtgtttactt	aagggtgatgt	cagtatttat	aggggtgtgga	gtgcaataaa	1500
atccgtgttg	actttaagtg	cgtgttttat	gactcagggg	tggggactgt	gggtatataa	1560
gcaggtgcag	acctgtgtgg	tcagttcaga	gcaggactca	tggagatctg	gactgtcttg	1620
gaagactttc	accagactag	acagttgcta	gagaactcat	cggagggagt	ctcttacctg	1680
tggagattct	gcttcgggtg	gcctctagct	aagctagtct	atagggccaa	acaggattat	1740
aaggaacaat	ttgaggatat	tttgagagag	tgtcctggta	tttttgactc	tctcaacttg	1800
ggccatcagt	ctcactttta	ccagagtatt	ctgagagccc	ttgacttttc	tactcctggc	1860
agaactaccg	ccgcggtagc	cttttttgcc	tttattcttg	acaaatggag	tcaagaaacc	1920
catttcagca	gggattaccg	tctggactgc	ttagcagtag	ctttgtggag	aacatggagg	1980
tgccagcgcc	tgaatgcaat	ctccggctac	ttgccagtag	agccggtaga	cacgctgagg	2040

atcctgagtc	tccagtcacc	ccaggaacac	caacgccgcc	agcagccgca	gcaggagcag	2100
cagcaagagg	aggaccgaga	agagaacccg	agagccggtc	tggaccctcc	ggtggcggag	2160
gaggaggagt	agctgacttg	tttcccagac	tgcgccgggt	gctgactagg	tcttccagtg	2220
gacgggagag	ggggattaag	cgggagaggc	atgaggagac	tagccacaga	actgaactga	2280
ctgtcagtct	gatgagccgc	aggcgcccag	aatcggtgtg	gtggcatgag	gtgcagtcgc	2340
aggggataga	tgaggtctcg	gtgatgcatg	agaaatattc	cctagaacaa	gtcaagactt	2400
gttggttga	gcccaggat	gattgggagg	tagccatcag	gaattatgcc	aagctggctc	2460
tgaagccaga	caagaagtac	aagattacca	aactgattaa	tatcagaaat	tcttcttaca	2520
tttcaggga	tggggccgag	gtggagatca	gtaccaggga	gaggggtggc	ttcagatggt	2580
gtatgatgaa	tatgtaccgc	ggggtgggtg	gcatggaggg	agtcaccttt	atgaacacga	2640
ggttcagggg	tgatgggtat	aatggggtgg	tctttatggc	caacaccaag	ctgacagtgc	2700
acggatgctc	cttctttggc	ttcaataaca	tgtgcatcga	ggcctggggc	agtgtttcag	2760
tgaggggatg	cagcttttca	gcgaactgga	tgggggtcgt	gggcagaacc	aagagcaagg	2820
tgtcagtga	gaaatgcctg	ttcgagaggt	gccacctggg	ggtgatgagc	gagggcggaag	2880
ccaaagtcaa	acactgcgcc	tctaccgaga	cgggctgctt	tgtgctgac	aagggcaatg	2940
cccaagtcaa	gcataacatg	atctgtgggg	cctcgcatga	gcgcggctac	cagatgctga	3000
cctgcgccgg	tgggaacagc	catatgctgg	ccaccgtgca	tgtggcctcg	caccccgca	3060
agacatggcc	cgagtccgag	cacaacgtca	tgaccgcgtg	caatgtgcac	ctgggctccc	3120
gccgaggcat	gttcatgccc	taccagtgca	acatgcaatt	tgtgaagggtg	ctgctggagc	3180
ccgatgccat	gtccagagtg	agcctgacgg	gggtgtttga	catgaatgtg	gagctgtgga	3240
aaattctgag	atatgatgaa	tccaagacca	ggtgccgggc	ctgcgaatgc	ggaggcaagc	3300
acgccaggct	tcagcccgtg	tgtgtggagg	tgacggagga	cctgcgaccc	gatcatttgg	3360
tgttgtcctg	caacgggacg	gagttcggct	ccagcgggga	agaatctgac	tagagtgagt	3420
agtgtttggg	gctgggtgtg	agcctgcatg	aggggcagaa	tgactaaaat	ctgtggtttt	3480
ctgtgtgttg	cagcagcatg	agcgggaagc	cctcctttga	gggaggggta	ttcagccctt	3540
atctgacggg	gcgtctcccc	tcttggggcg	gagtgcgtca	gaatgtgatg	ggatccacgg	3600
tggacggccg	gcccgtgcag	cgcgcgaact	cttcaaccct	gacctacgcg	accctgagct	3660
cctcgtccgt	ggacgcagct	gccgcgcgag	ctgctgcttc	cgccgccagc	gccgtgcgcg	3720
gaatggccct	gggcgccggc	tactacagct	ctctggtggc	caactcgagt	tccaccaata	3780

atccccccag cctgaacgag gagaagctgc tgctgctgat ggcccagctc gaggccctga	3840
cccagcgcct gggcgagctg acccagcagg tggtcagct gcaggcggag acgcgggccc	3900
cggttgccac ggtgaaaacc aaataaaaaa tgaatcaata aataaacgga gacggttggt	3960
gattttaaca cagagtcttg aatcttttatt tgatttttcg cgegcggtag gccctggacc	4020
accggtctcg atcattgagc acccgggtgga tcttttccag gacccggtag aggtgggctt	4080
ggatgttgag gtacatgggc atgagcccgt cccgggggtg gaggtagctc cattgcaggg	4140
cctcgtgctc ggggatgggtg ttgtaaatca cccagtcata gcaggggcgc agggcggtgt	4200
gctgcacgat gtccttgagg aggagactga tggccacggg cagccccttg gtgtaggtgt	4260
tgacgaacct gttgagctgg gagggatgca tgcgggggga gatgagatgc atcttggcct	4320
ggatcttgag attggcgatg tccccgccc gatcccgccg ggggttcattg ttgtgcagga	4380
ccaccagcac ggtgtatccg gtgcacttgg ggaatttgtc atgcaacttg gaaggaagg	4440
cgtgaaagaa tttggagacg cccttgtagc cggccagggtt ttccatgcac tcatccatga	4500
tgatggcgat gggcccgtgg gcggcgccct gggcaaaagac gtttcggggg tcggacacat	4560
cgtagttgtg gtcctgggtg agctcgtcat aggccatttt aatgaatttg gggcgagggg	4620
tgcccgactg ggggacgaag gtgccctcga tcccgggggc gtagttgcc tcgcagatct	4680
gcatctccca ggccttgagc tcggaggggg ggatcatgtc cacctgcggg gcgatgaaaa	4740
aaacggtttc cggggcgggg gagatgagct gggccgaaag caggttccgg agcagctggg	4800
acttgccga accggtgggg ccgtagatga ccccgatgac cggctgcagg tggtagttga	4860
gggagagaca gctgccgtcc tcgcggagga gggggggccac ctcgttcatc atctcgcgca	4920
catgcatgtt ctgcgcacg agttccgcca ggaggcgctc gccccccagc gagaggagct	4980
cttgacgca ggcgaagttt ttcagcggct tgagtcgctc ggccatgggc attttggaga	5040
gggtctgttg caagagttcc agacgggtccc agagctcggg gatgtgctct agggcatctc	5100
gatccagcag acctcctcgt ttgcggggtt ggggcgactg cgggagtagg gcaccaggcg	5160
atgggcgtcc agcaggcca gggtcgggtc cttccagggc cgcagggtcc gcgtcagcgt	5220
ggctccgctc acggtgaagg ggtgcgcgcc gggctgggcg cttgcgaggg tgcgttcag	5280
gctcatccgg ctggtcgaga accgctccc gtcggcgccc tgcgcgtcgg ccaggtagca	5340
attgagcatg agttcgtagt tgagcgctc ggccgcgtgg cccttggcgc ggagcttacc	5400
tttggagtg tgcgcgaga cgggacagag gagggacttg agggcgtaga gcttgggggc	5460

gaggaagacg gactcggggg cgtaggcgtc cgcgccgcag ctggcgcaga cggctctcgca	5520
ctccacgagc caggtgaggt cggggcggtt ggggtcaaaa acgaggtttc ctccgtgctt	5580
tttgatgcgt ttcttacctc tgggtctccat gagctcgtgt ccccgctggg tgacaaagag	5640
gctgtccgtg tcccgcgtaga ccgactttat gggccgggtcc tcgagcgggg tgccgcgggtc	5700
ctcgtcgtag aggaaccccg cccactccga gacgaaggcc cgggtccagg ccagcacgaa	5760
ggaggccacg tgggaggggt agcggtcgtt gtccaccagc ggggtccacct tctccaggg	5820
atgcaagcac atgtccccct cgtccacatc caggaagggt attggcttgt aagtgtaggc	5880
cacgtgaccg ggggtcccg cgggggggt ataaaagggg gcggggccct gctcgtcctc	5940
actgtcttcc ggatcgctgt ccaggagcgc cagctgttgg ggtaggtatt ccctctcgaa	6000
ggcgggcatg acctcggcac tcaggttgtc agtttctaga aacgaggagg atttgatatt	6060
gacgggtgccg ttggagacgc ctttcatgag cccctcgtcc atttggtcag aaaagacgat	6120
ctttttgttg tcgagcttgg tggcgaagga gccgtagagg gcgttggaga gcagcttggc	6180
gatggagcgc atggtctggt tcttttcctt gtccggcgcgc tccttggcgg cgatgttgag	6240
ctgcacgtac tcgcgcgcca cgcacttcca ttccgggaag acggtggtga gctcgtcggg	6300
cacgattctg acccgccagc cgcggttgtg cagggtagt aggtccacgc tgggtgccac	6360
ctcgccgcgc aggggctcgt tgggtccagca gaggcgccc ccttgccgcg agcagaagg	6420
gggcagcggg tccagcatga gctcgtcggg ggggtcggcg tccacggtga agatgccggg	6480
caggagctcg gggtcgaagt agctgatgca ggtgccaga ttgtccagcg ccgcttgcca	6540
gtcgcgcacg gccagcgcgc gctcgtaggg gctgaggggc gtgcccagg gcattggggtg	6600
cgtgagcgcg gaggcgtaca tgcgcgagat gtcgtagac tagaggggct cctcgaggac	6660
gccgatgtag gtggggtagc agcgcctccc gcggatgctg gcgcgcacgt agtcgtacag	6720
ctcgtgcgag ggcgcgagga gcccgtgcc gaggttggag cgttgccgct tttcggcgcg	6780
gtagacgac tggcggaaga tggcgtggga gttggaggag atgggtgggccc tttggaagat	6840
gttgaagtgg gcgtggggca ggccgaccga gtccctgatg aagtgggcgt aggaatcctg	6900
cagcttggcg acgagctcgg cggtgacgag gacgtccagg gcgcagtagt cgagggtctc	6960
ttggatgatg tcatacttga gctggccctt ctgcttcac agctcgcggg tgagaaggaa	7020
ctcttcgcgg tccttccagt actcttcgag ggggaacccg tcctgatcgg cacggttaaga	7080
gcccaccatg tagaactggt tgacggcctt gtaggcgcag cagcccttct ccacggggag	7140
ggcgtaagct tgccgggcct tgccgagggg ggtgtgggtg agggcggaagg tgcgcgcac	7200

catgaccttg aggaactggt gcttgaagtc gaggtcgtcg cagccgccct gctcccagag	7260
ttggaagtcc gtgcgcttct tgtaggcggg gttaggcaaa gcgaaagtaa catcgttgaa	7320
gaggatcttg cccgcgcggg gcatgaagtt gcgagtgatg cggaaaggct ggggcacctc	7380
ggcccgggtg ttgatgacct gggcggcgag gacgatctcg tcgaagccgt tgatgttgtg	7440
cccgacgatg tagagttcca cgaatcgcgg gcggcccttg acgtggggca gcttcttgag	7500
ctcgtcgtag gtgagctcgg cggggtcgct gagcccgctg tgctcgaggg ccagtcggc	7560
gacgtggggg ttggcgctga ggaaggaggt ccagagatcc acggccaggg cggctctgcaa	7620
gcgggtcccgg tactgacgga actgttgggc cacggccatt ttttcggggg tgacgcagta	7680
gaagggtgcgg gggtcgccgt gccagcggtc ccacttgagc tggagggcga ggtcgtgggc	7740
gagctcgacg agcggcgggt ccccgagagag tttcatgacc agcatgaagg ggacgagctg	7800
cttgccgaag gaccccatcc aggtgtaggt ttccacatcg taggtgagga agagcctttc	7860
ggtgcgagga tgcgagccga tggggaagaa ctggatctcc tgccaccagt tggaggaatg	7920
gctgttgatg tgatggaagt agaaatgccg acggcgcgcc gagcactcgt gcttgtgttt	7980
atacaagcgt ccgcagtgtc cgcaacgctg cacgggatgc acgtgtgca cgagctgtac	8040
ctgggttcct ttggcgagga atttcagtgg gcagtggagc gctggcggct gcctctcgtg	8100
ctgtactacg tcttggccat cggcgtggcc atcgtctgcc tcgatggtgg tcatgctgac	8160
gagcccgcg cggaggcagg tccagacctc ggctcggacg ggtcggagag cgaggacgag	8220
ggcgcgagg ccggagctgt ccagggtcct gagacgctgc ggagtcaggt cagtgggcag	8280
cggcgcgcg cggttgactt gcaggagctt ttccagggcg cgcgggaggt ccagatggta	8340
cttgatctcc acggcgccgt tggtaggtac gtccacggct tgcagggtgc cgtgcccctg	8400
gggcgccacc accgtgcccc gtttcttctt gggcgctgct tccatgtcgg tcagaagcgg	8460
cggcgaggac gcgcgccggg cggcaggggc ggctcggggc ccggaggcag gggcggcagg	8520
ggcacgtcgg cgcgcgcgc gggcagggtt tggtagtcg cccggagaag actggcgtga	8580
gcgacgacgc gacggttgac gtcctggatc tgacgcctct gggatgaagg cacgggaccc	8640
gtgagtttga acctgaaaga gaggtcgaca gaatcaatct cggatcgtt gacggcggcc	8700
tgccgcagga tctcttgac gtcgccgag ttgtcctggt aggcgatctc ggtcatgaac	8760
tgctcgatct cctcctctg aaggctctcg cggccggcgc gctcgacggt ggcgcgagg	8820
tcgttgagga tgcggcccat gagctgcgag aaggcgttca tgccggcctc gttccagacg	8880

cggtctgtaga	ccacgggtcc	gtcgggggtcg	cgcgcgcgca	tgaccacctg	ggcgaggttg	8940
agctcgacgt	ggcgcgtaga	gaccggttag	ttgcagaggc	gctggtagag	gtagttgagc	9000
gtggtggcga	tgtgctcggt	gacgaagaag	tacatgatcc	agcggcgagg	cggcatctcg	9060
ctgacgtcgc	ccagggcttc	caagcggttc	atggcctcgt	agaagtcac	ggcgaagttg	9120
aaaaactggg	agttgcgcgc	cgagacggtc	aactcctcct	ccagaagacg	gatgagctcg	9180
gcgatgggtg	cgcgcacctc	gcgctcgaag	gccccggggg	gctcctcttc	catctcctcc	9240
tcttctcctc	ccactaacat	ctcttctact	tcctcctcag	gaggcggttg	cgggggaggg	9300
gccctgcgtc	gccggcgggc	cacgggcaga	cggtcgatga	agcgctcgat	ggtctccccg	9360
cgcggcgac	gcatgggttc	ggtgacggcg	cgcccgctct	cgcggggccg	cagcatgaag	9420
acgcgcgcgc	gcatctccag	gtggccggcg	gggggggtctc	cggtgggcag	ggagagggcg	9480
ctgacgatgc	atcttatcaa	ttgaccgcta	gggactccgc	gcaaggacct	gagcgtctcg	9540
agatccacgg	gatccgaaaa	ccgctgaacg	aaggcttcga	gccagtcgca	gtcgcaaggt	9600
aggctgagcc	cggtttcttg	ttcttcgggt	atttggtcgg	gaggcgggcg	ggcgatgctg	9660
ctggtgatga	agttgaagta	ggcggtcctg	agacggcgga	tggtaggcag	gagcaccagg	9720
tccttggggc	cggcttgctg	gatgcgcgca	cggtcgggca	tgccccaggc	gtggctctga	9780
cacctggcga	ggtccttgta	gtagtcctgc	atgagccgct	ccacgggcac	ctcctcctcg	9840
cccgcgcggc	cgtgcatcg	cgtgagcccc	aaccgcgcgt	gcggctggac	gagcgccagg	9900
tcggcgacga	cgcgctcggt	gaggatggcc	tgctggatct	gggtgagggg	ggtctggaag	9960
tcgtcgaagt	cgacgaagcg	gtggtaggct	ccgggtgtga	tgggttagga	gcagttggcc	10020
atgacggacc	agttgacggg	ctgggtggcg	ggtcgcacga	gctcgtggta	cttgaggcgc	10080
gagtaggcgc	gcgtgtcgaa	gatgtagtcg	ttgcaggcgc	gcacgaggta	ctggatatccg	10140
acgaggaagt	gcggcgggcg	ctggcggttag	agcgggccatc	gctcgggtggc	ggggggcgccg	10200
ggcgcgaggt	cctcgagcat	gaggcggtgg	tagccgtaga	tgtacctgga	catccagggtg	10260
atgccggcg	cgggtggtgga	ggcgcgcggg	aactcgcgga	cgcggttcca	gatgttgcg	10320
agcggcagga	agtagttcat	ggtggcgcg	gtctggcccc	tgaggcgcg	gcagtcgtgg	10380
atgctctaga	catacgggca	aaaacgaaag	cggtcagcgg	ctcgactccg	tggcctggag	10440
gctaagcgaa	cgggttgggc	tgcgcggtga	ccccgggttcg	aatctcgaat	caggctggag	10500
ccgcagctaa	cgtggtactg	gcactcccgt	ctcgacccaa	gcctgctaac	gaaacctcca	10560
ggatacggag	gcgggtcggt	ttttggcctt	ggtcgctggg	catgaaaaac	tagtaagcgc	10620

ggaaagcggc	cgcccgcgat	ggctcgcctgc	cgtagtctgg	agaaagaatc	gccaggggtg	10680
cgttgcggtg	tgcccgggtt	cgagcctcag	cgctcggcgc	cggccggatt	ccgcggctaa	10740
cgtgggctg	gctgccccgt	cgtttccaag	accccttagc	cagccgactt	ctccagttac	10800
ggagcgagcc	cctctttttt	tttcttgtgt	ttttgcgaga	tgcattccgt	actgcggcag	10860
atgcgcccc	accctccacc	acaaccgccc	ctaccgcagc	agcagcaaca	gccggcgctt	10920
ctgccccgc	cccagcagca	gccagccact	accgcggcgg	ccgcctgag	cggagccggc	10980
gttcagtatg	acctggcctt	ggaagagggc	gaggggctgg	cgcggctggg	ggcgtcgtcg	11040
ccggagcggc	accgcgcgt	gcagatgaaa	agggacgctc	gcgaggccta	cgtgccccag	11100
cagaacctgt	tcagagacag	gagcggcgag	gagcccgagg	agatgcgcgc	ctcccgcttc	11160
cacgcggggc	gggagctgcg	gcgcggcctg	gaccgaaagc	gggtgctgag	ggacgaggat	11220
ttcgaggcgg	acgagctgac	ggggatcagc	cccgcgcgcg	cgcacgtggc	cgcggccaac	11280
ctgggtcacg	cgtacgagca	gaccgtgaag	gaggagagca	acttccaaaa	atccttcaac	11340
aaccacgtgc	gcacgctgat	cgcgcgcgag	gaggtgacct	tgggcctgat	gcacctgtgg	11400
gacctgctgg	aggccatcgt	gcagaacccc	acgagcaagc	cgctgacggc	gcagctgttt	11460
ctgggtggtg	agcacagtgc	ggacaacgag	acgttcaggg	aggcgctgct	gaatatcacc	11520
gagcccgagg	gccgctggct	cctggacctg	gtgaacattt	tgcagagcat	cgtggtgcag	11580
gagcgcgggc	tgccgctgtc	cgagaagctg	gcggccatca	acttctcggt	gctgagtctg	11640
ggcaagtact	acgctaggaa	gatctacaag	accccgtagc	tgcccataga	caaggagggtg	11700
aagatcgacg	ggttttacat	gcgcgatgacc	ctgaaagtgc	tgacctgag	cgacgatctg	11760
ggggtgtacc	gcaacgacag	gatgcaccgc	gcggtgagcg	ccagccgcgg	gcgcgagctg	11820
agcgaccagg	agctgatgca	cagcctgcag	cgggccctga	ccggggccgg	gaccgagggg	11880
gagagctact	ttgacatggg	cgcggacctg	cgctggcagc	ccagccgcgg	ggccttggaa	11940
gctgccggcg	gttcccccta	cgtggaggag	gtggacgatg	aggaggagga	gggcgagtac	12000
ctggaagact	gatggcgcca	ccgtattttt	gctagatgca	gcaacagcca	ccgcgcgcgc	12060
ctcctgatcc	cgcgatgcgg	gcggcgctgc	agagccagcc	gtccggcatt	aactcctcgg	12120
acgattggac	ccaggccatg	caacgcacga	tggcgctgac	gacccgcaat	cccgaagcct	12180
ttagacagca	gcctcaggcc	aaccggctct	cggccatcct	ggaggccgtg	gtgccctcgc	12240
gctcgaaccc	cacgcacgag	aagggtgctg	ccatcgtgaa	cgcgctggtg	gagaacaagg	12300

ccatccgcgg tgacgaggcc gggctggtgt acaacgcgct gctggagcgc gtggcccgcct 12360
acaacagcac caacgtgcag acgaacctgg accgcatggt gaccgacgtg cgcgaggcgg 12420
tgtcgcagcg cgagcgggtc caccgcgagt cgaacctggg ctccatgggt gcgctgaacg 12480
ccttcctgag cagcagccc gccaacgtgc cccggggcca ggaggactac accaacttca 12540
tcagcgcgct gcggctgatg gtggcgcagg tgcccagag cgagggtgtac cagtgcggggc 12600
cggactactt cttccagacc agtcgccagg gcttgacagc cgtgaacctg agccaggctt 12660
tcaagaactt gcagggactg tggggcgctg agggcccggg cggggaccgc gcgacggtgt 12720
cgagcctgct gacgccgaac tcgcgcctgc tgctgctgct ggtggcgccc ttacaggaca 12780
gcggcagcgt gagccgcgac tcgtacctgg gctacctgct taacctgtac cgcgaggcca 12840
tcggacaggg gcacgtggac gagcagacct accaggagat caccacgtg agccgcgcgc 12900
tgggccagga ggaccggggc aacctggagg ccacctgaa cttcctgctg accaaccggg 12960
cgcagaagat cccgccccag tacgcgctga gcaccgagga ggagcgcac ctgcgctacg 13020
tgacgacag cgtaggggctg ttcttgatgc aggagggggc cagccccagc gcggcgctcg 13080
acatgaccgc gcgcaacatg gagcccagca tgtacgcccg caaccgcccg ttcacataa 13140
agctgatgga ctacttgcat cggggcgccg ccatgaactc ggactacttt accaagcca 13200
tcttgaacct gcactggctc ccgcccccg ggttctacac gggcgagtag gacatgcccg 13260
accccaacga cgggttcctg tgggacgacg tggacagcag cgtgttctcg ccgcgtccag 13320
gaaccaatgc cgtgtggaag aaagagggcg gggaccggcg gccgtcctcg gcgctgtccg 13380
gtcgcgcggg tgctgcccg gcggtgcccg agggcccgag ccccttcccg agcctgccct 13440
tttcgctgaa cagcgtgcgc agcagcgagc tgggtcggct gacgcgaccg cgcctgctgg 13500
gcgaggagga gtacctgaac gactccttgt tgaggcccg gcgcgagaag aacttcccca 13560
ataacgggat agagagcctg gtggacaaga tgagccgctg gaagacgtac ggcacagagc 13620
acagggacga gccccgagct agcagcgag gcacccgtag acgccagcgg cagcagaggc 13680
agcggggact ggtgtgggac gatgaggatt ccgccgacga cagcagcgtg ttggacttgg 13740
gtgggagtg tggttaaccg ttccgtcacc tcgcccccg tatcgggcgc ctgatgtaag 13800
aatctgaaaa aataaaagac ggtactcacc aaggccatgg cgaccagcgt gcgttcttct 13860
ctgttggttg tagtagtatg atgaggcgcg tgtacccgga gggtcctcct cctcgtacg 13920
agagcgtgat gcagcaggcg gtggcggcg cgatgcagcc cccgctggag gcgccttacg 13980
tgcccccgcg gtacctggcg cctacggagg ggcggaacag cattcgttac tcggagctgg 14040

cacccttgta cgataccacc cggttgtacc tgggtggacaa caagtcggca gacatcgctt 14100
cgctgaacta ccagaacgac cacagcaact tcctgaccac cgtggtgcag aacaacgatt 14160
tcacccccac ggaggccagc acccagacca tcaactttga cgagcgctcg cgggtggggcg 14220
gccagctgaa aaccatcatg cacaccaaca tgcccaacgt gaacgagttc atgtacagca 14280
acaagttcaa ggcgcgggtg atggtctcgc gcaagacccc caacgggggtg gatgatgatt 14340
atgatggtag tcaggacgag ctgacctacg agtgggtgga gtttgagctg cccgagggca 14400
acttctcggg gaccatgacc atcgatctga tgaacaacgc catcatcgac aactacttgg 14460
cgggtggggcg gcagaacggg gtgctggaga gcgacatcgg cgtgaagttc gacacgcgca 14520
acttcgggct gggctgggac cccgtgaccg agctggtgat gccgggctg tacaccaacg 14580
aggccttcca ccccgacatc gtccctgctgc ccggtgcgg cgtggacttc accgagagcc 14640
gcctcagcaa cctgctgggc atccgcaagc ggagccctt ccaggagggc ttccagatcc 14700
tgtacgagga cctggagggg ggcâacatcc ccgcgctctt ggatgtcgaa gcctacgaga 14760
aaagcaagga ggatagcacc gccgcggcga ccgcagccgt ggccaccgcc tctaccgagg 14820
tgcggggcca taattttgct agcgtgcgg cagcgccga ggcggctgaa accgaaagta 14880
agatagtcac ccagccgggtg gagaaggaca gcaaggacag gagctacaac gtgctcgcgg 14940
acaagaaaaa caccgcctac cgcagctggt acctggccta caactacggc gaccccgaga 15000
agggcgctgc ctctggacg ctgctacca cctcggacgt cacctgcggc gtggagcaag 15060
tctactggtc gctgcccgac atgatgcaag acccggtcac ctccgctcc acgcgtcaag 15120
ttagcaacta cccggtgggt ggcgcgagc tcctgcccggt ctactccaag agcttcttca 15180
acgagcaggc cgtctactcg cagcagctgc gcgccttcac ctgcgtcacg cacgtcttca 15240
accgttccc cgagaaccag atcctcgtcc gccgcgccgc gccaccatt accaccgtca 15300
gtgaaaacgt tcctgctctc acagatcacg ggaccctgcc gctgcgcagc agtatccggg 15360
gagtcacgag cgtgaccgtc actgacgcca gacgcgcac ctgcccctac gtctacaagg 15420
ccctgggctg agtcgcgcgc cgcgtctctc cgagccgcac cttctaaaaa atgtccattc 15480
tcctctcgcc cagtaataac accggttggg gcctgcgcgc gccagcaag atgtacggag 15540
gcgctcgcca acgctccacg caacaccccg tgcgctgcgc cgggcacttc cgcgtccct 15600
ggggcgccct caagggccgc gtgcgctcgc gcaccaccgt cgacgacgtg atcgaccagg 15660
tgggtggccga cgcgcgcaac tacacgcccg ccgccgcgc cgtctccacc gtggacgccc 15720

tcatcgacag	cgtggtggcc	gacgcgcgcc	ggtacgccc	caccaagagc	cggcggcgcc	15780
gcatcgccc	gcggcaccgg	agcaccccc	ccatgcgcgc	ggcgcgagcc	ttgctgcgca	15840
gggccaggcg	cacgggacgc	agggccatgc	tcaggcgcc	cagacgcgcg	gcctccggca	15900
gcagcagcg	cggcaggacc	cgcagacgcg	cggccacggc	ggcgcgcgcg	gccatcgcca	15960
gcatgtccc	cccgcggcgc	ggcaacgtgt	actgggtgcg	cgacgcccgc	accgggtgtgc	16020
gcgtgcccgt	gcgcacccgc	ccccctcgca	cttgaagatg	ctgacttcgc	gatgttgatg	16080
tgtcccagcg	gcgaggagga	tgtccaagcg	caaatacaag	gaagagatgc	tccaggatcat	16140
cgcgcctgag	atctacggcc	ccgcggcgcc	ggtgaaggag	gaaagaaagc	cccgcgaaact	16200
gaagcgggtc	aaaaaggaca	aaaaggagga	ggaagatgac	ggactggtgg	agtttgtgcg	16260
cgagttcgcc	ccccggcgcc	gcgtgcagtg	gcgcggcgcg	aaagtgaaac	cgggtgctgcg	16320
gcccggcacc	acggtggtct	tcacgcccgg	cgagcggtcc	ggctccgcct	ccaagcgctc	16380
ctacgacgag	gtgtacgggg	acgaggacat	cctcgagcag	gcggtcgagc	gtctgggcca	16440
gtttgcgtac	ggcaagcgca	gccgccccgc	gcccttgaaa	gaggaggcgcg	tgtccatccc	16500
gctggaccac	ggcaacccca	cgcgcgagcct	gaagccggtg	accctgcagc	aggtgctacc	16560
gagcgcggcg	ccgcgcgggg	gcttcaagcg	cgaggcgcc	gaggatctgt	acccgaccat	16620
gcagctgatg	gtgcccgaagc	gccagaagct	ggaggacgtg	ctggagcaca	tgaagggtgga	16680
ccccgagggtg	cagcccgagg	tcaagggtgcg	gcccatacaag	cagggtggccc	cgggcctggg	16740
cgtgcagacc	gtggacatca	agatccccac	ggagcccatg	gaaacgcaga	ccgagcccgt	16800
gaagcccagc	accagcacca	tggagggtgca	gacggatccc	tggatgccag	caccagcttc	16860
caccagcact	cgccgaagac	gdaagtacgg	cgcggccagc	ctgctgatgc	ccaactacgc	16920
gctgcaccc	tccatcatcc	ccacgcgggg	ctaccgcggc	acgcgcttct	accgcggcta	16980
caccagcagc	cgcgcggcca	agaccaccac	ccgcgcgggt	cgtcgagccc	gccgcagcag	17040
caccgcgact	tccgccttgg	tgcggagagt	gtatcgagc	ggcgcgagc	ctctgaccct	17100
gccgcgcgcg	cgctaccacc	cgagcatcgc	catttaacta	ccgcctccta	cttgagata	17160
tggccctcac	atgccgcctc	cgcgtcccca	ttacgggcta	ccgaggaaga	aagccgcgcc	17220
gtagaaggct	gacgggggaa	gggctgcgtc	gccatcacca	ccggcgcgcg	cgcgccatca	17280
gcaagcgggt	ggggggaggc	ttcctgccc	cgctgatccc	catcatcgcc	gcggcgatcg	17340
gggcgatccc	cggcatagct	tccgtggcg	tgcaggcctc	tcagcgccac	tgagacacaa	17400
aaaagcatgg	atttgtaata	aaaaaaaaa	tggactgacg	ctcctggtcc	tgtgatgtgt	17460

gtttttagat ggaagacatc aatttttcgt ccctggcacc gcgacacggc acgcgggcgt 17520
 ttatgggcac ctggagcgac atcggaaca gccaaactgaa cgggggccc ttcaattgga 17580
 gcagtctctg gagcgggctt aagaatttcg ggtccacgct caaaacctat ggcaacaagg 17640
 cgtggaacag cagcacaggg caggcgctga gggaaaagct gaaagaacag aacttccagc 17700
 agaaggtggt tgatggcctg gcctcaggca tcaacggggt ggttgacctg gccaaaccagg 17760
 ccgtgcagaa acagatcaac agccgcctgg acgcggtccc gcccgcgggg tccgtggaga 17820
 tgccccaggt ggaggaggag ctgcctcccc tggacaagcg cggcgacaag cgaccgcgtc 17880
 ccgacgcgga ggagacgctg ctgacgcaca cggacgagcc gccccgtac gaggaggcgg 17940
 tgaaactggg cctgcccacc acgcgggccc tggcgctct ggcaccgga gtgctgaaac 18000
 ccagcagcag ccagcccgcg accctggact tgcctccgcc tcgccccctcc acagtggcta 18060
 agccccctgcc gccggtggcc gtgcgctgc gcgccccccg aggcgcgccc caggcgaact 18120
 ggcagagcac tctgaacagc atcgtgggtc tgggagtgc gagtgtaag cgccgccgct 18180
 gctattaaaa gacactgtag cgcttaactt gcttgtctgt gtgtatatgt atgtccgccg 18240
 accagaagga ggagtgtgaa gaggcgctc gccgagttgc aagatggcca ccccatcgat 18300
 gctgccccag tgggcgtaca tgcacatcgc cggacaggac gcttcggagt acctgagtc 18360
 ggggtctggtg cagttcgccc gcgccacaga cacctacttc agtctgggga acaagtttag 18420
 gaacccccacg gtggcgccca cgcacgatgt gaccaccgac cgcagccagc ggctgacgct 18480
 gcgcttcgtg cccgtggacc gcgaggacaa cacctactcg taaaaagtgc gctacacgct 18540
 ggccgtgggc gacaaccgcg tgctggacat ggccagcacc tactttgaca tccgcggcgt 18600
 gctggaccgg ggccctagct tcaaacccta ctctggcacc gcctacaaca gcctagctcc 18660
 caagggagct cccaattcca gccagtggga gcaagcaaaa acaggcaatg ggggaactat 18720
 ggaaacacac acatatggtg tggccccaat gggcgagag aatattacaa aagatggtct 18780
 tcaaattgga actgacgtta cagcgaatca gaataaacca atttatgccg aaaaaacatt 18840
 tcaaccagaa ccgcaagtag gagaagaaaa ttggcaagaa actgaaaact tttatggcgg 18900
 tagagctctt aaaaaagaca caaacatgaa accttgctat ggctcctatg ctagaccac 18960
 caatgaaaaa ggaggtcaag ctaaaactta agttggagat gatggagttc caaccaaaga 19020
 attcgacata gacctggctt tctttgatac tcccggtggc accgtgaacg gtcaagacga 19080
 gtataaagca gacattgtca tgtatáccga aaacacgtat ttggaaactc cagacacgca 19140

tgtggtatac	aaaccaggca	aggatgatgc	aagttctgaa	attaacctgg	ttcagcagtc	19200
tatgccccaac	agacccaact	acattggggt	cagggacaac	tttatcggtc	ttatgtacta	19260
caacagcact	ggcaatatgg	gtgtgcttgc	tggtcaggcc	tcccagctga	atgctgtggg	19320
tgatttgcaa	gacagaaaca	ccgagctgtc	ctaccagctc	ttgcttgact	ctttgggtga	19380
cagaacccgg	tatttcagta	tgtggaacca	ggcgggtggac	agttatgacc	ccgatgtgcg	19440
catcatcgaa	aaccatgggt	tggaggatga	attgccaaac	tattgcttcc	ccttgggacgg	19500
ctctggcact	aacgccgcat	accaagggtg	gaaagtaaaa	gatgggtcaag	atgggtgatgt	19560
tgagagtga	tgggaaaatg	acgatactgt	tgcagctcga	aatcaattat	gtaaaggtaa	19620
cattttcgcc	atggagatta	atctccaggc	taacctgtgg	agaagtttcc	tctactcgaa	19680
cgtggccctg	tacctgcccg	actcctacaa	gtacacgccg	accaacgtca	cgtgcccagc	19740
caacaccaac	acctacgatt	acatgaatgg	cagagtgaac	cctccctcgc	tggtagacgc	19800
ctacctcaac	atcggggcgc	gctgggtcgt	ggaccccatg	gacaacgtca	acccttcaa	19860
ccaccaccgc	aacgcggggc	tgcgtaccgc	ctccatgctc	ctgggcaacg	ggcgctacgt	19920
gcccttccac	atccagggtc	cccaaaagtt	tttcgccatc	aagagcctcc	tgctcctgcc	19980
cgggtcctac	acctacgagt	ggaacttccg	caaggacgtc	aacatgatcc	tgacagagctc	20040
cctaggcaac	gacctgcgca	cggacggggc	ctccatcgcc	ttcacagca	tcaacctcta	20100
cgccaccttc	ttccccatgg	cgcacaacac	cgcctccacg	ctcgaggcca	tgctgcgcaa	20160
cgaaccaaac	gaccagtcct	tcaacgacta	cctctcggcg	gccaacatgc	tctaccccat	20220
cccggccaac	gccaccaacg	tgcccatctc	catcccctcg	cgaactggg	ccgccttccg	20280
cggatggtec	ttcacgcgcc	tgaagacceg	cgaagacgcc	tcgctcggct	ccgggttcga	20340
cccctacttc	gtctactcgg	gtccatccc	ctacctagac	ggcaccttct	acctcaacca	20400
caccttcaag	aagggtctcca	tcaccttcga	ctcctccgtc	agctggcccg	gcaacgaccg	20460
cctcctgacg	cccacgagt	tcaaaatcaa	gcgcaccgtc	gacggagagg	gatacaacgt	20520
ggcccagtg	aacatgacca	aggactgggt	cctgggtccag	atgctggccc	actacaacat	20580
cggctaccag	ggcttctacg	tgcccagggg	ctacaaggac	cgcagtact	ccttcttccg	20640
caacttccag	cccatgagcc	gccagggtcg	ggacgaggtc	aactacaagg	actacagggc	20700
cgtcaccctg	gcctaccagc	acaacaactc	gggcttcgtc	ggctacctcg	cgcaccat	20760
gcgccagggc	cagccctacc	ccgccaaacta	cccctaccgg	ctcatcgga	agagcgccgt	20820
cgcacagcgc	accagaaaa	agttcctctg	cgaacgggtc	atgtggcgca	tccccttctc	20880

cagcaacttc atgtccatgg gcgcgctcac cgacctcggc cagaacatgc tctacgccaa 20940
ctccgcccac gcgctagaca tgaatttcga agtcgacccc atggatgagt ccaccttct 21000
ctatgttgtc ttcgaagtct tcgacgtcgt ccgagtgcac cagccccacc gggcgctcat 21060
cgaagccgtc tacctgcgca cggccttctc ggccggcaac gccaccacct aagccgctct 21120
tgcttcttgc aagatgacgg cgggctcggc cgagcaggag ctcagggccca tcctccgga 21180
cctgggctgc gggccctgct tcctgggcac cttcgacaag cgcttccctg gattcatggc 21240
cccgcacaag ctggcctgcg ccacgtgaa cacggccggc cgcgagaccg gggcgagca 21300
ctggctggcc ttcgcctgga acccgcgtc ccacacatgc tacctcttcg accccttcgg 21360
gttctcggac gagcgctca agcagatcta ccagttcgag tacgagggcc tgctgcgtcg 21420
cagcgccctg gccaccgagg accgctgct caccctggaa aagtcaccc agaccgtgca 21480
gggtccggc tcggccgctt ggggctctt ctgctgcatg ttctgcacg ccttcgtgca 21540
ctggcccgac cggcccatgg acaagaacc caccatgaac ttactgacgg gggcgccaa 21600
cggcatgctc cagtcgcccc aggtggaacc caccctgcgc cgcaaccagg aagcgtctta 21660
ccgcttctc aatgccact ccgcctactt tcgctccac cgcgcgcga tcgagaaggc 21720
caccgccttc gaccgcatga atcaagacat gtaaaaaacc ggtgtgtgta tgtaatgct 21780
ttattcataa taaacagcac atgtttatgc caccctctct gaggtctga cttatttag 21840
aaatcgaagg ggttctgccc gctctggca tggcccgcg gcagggatac gttgcggaac 21900
tggtacttgg gcagccactt gaactcgggg atcagcagct tgggcacggg gaggtcgggg 21960
aacgagtcgc tccacagctt gcgcgtgagt tgcaggcgcc ccagcaggtc gggcgcgag 22020
atcttgaaat cgcagttggg acccgcgttc tgcgcgcgag agttgcggta cagggggtg 22080
cagcactgga acaccatcag ggcggggtgc ttcacgcttg ccagcaccgt cgcgtcggtg 22140
atgccctcca cgtccagatc ctggcggtt gccatccga agggggtcac cttgcaggtc 22200
tgccgccccca tgctgggcac gcagccgggc ttgtggttg aatcgcagtg cagggggatc 22260
agcatcatct gggcctgctc ggagctcatg cccgggtaca tggccttcac gaaagcctcc 22320
agctggcgga aggcctgctg cgccttgccg ccctcgggtga agaagacccc gcaggacttg 22380
ctagagaact ggttggtggc gcagcggcg tcgtgcacgc agcagcgcg gtcgttggtg 22440
gccagctgca ccacgtgcg ccccagcgg ttctgggtga tcttgccccg gttgggggtc 22500
tccttcagcg cgcgtgccc gttctcgtc gccacatcca tctcgatagt gtgctccttc 22560

tggatcatca	cgggtcccggtg	caggcaccgc	agcttgcctt	cggcttcggt	gcagccgtgc	22620
agccacagcg	cgcagccggt	gcactcccag	ttcttgtggg	cgatctggga	gtgcgagtgc	22680
acgaagccct	gcaggaagcg	gcccattcatc	gcggtcaggg	tcttgttgct	ggtgaaggtc	22740
agcgggatgc	cgcggtgctc	ctcgttcaca	tacagggtggc	agatgcggcg	gtacacctcg	22800
ccctgctcgg	gcatcagctg	gaaggcggac	ttcagggtcgc	tctccacgcg	gtaccgggtcc	22860
atcagcagcg	tcattcacttc	catgcccttc	tcccaggccg	aaacgatcgg	caggctcagg	22920
gggttcttca	cgcctattgt	catcttagtc	gccgcgcgcg	aggtcagggg	gtcgttctcg	22980
tccaggggtct	caaacactcg	cttgccgtcc	ttctcgatga	tgcgacaggg	gggaaagctg	23040
aagcccacgg	cgcgcagctc	ctcctcggcc	tgcctttcgt	cctcgctgtc	ctggctgatg	23100
tcttgcaaag	gcacatgctt	ggtcttgcg	ggtttctttt	tgggcggcag	aggcggcggc	23160
gatgtgctgg	gagagcgga	gttctcgctc	accacgacta	tttcttcttc	ttggccgctc	23220
tccgagacca	cgcggcggtg	ggcatgcctc	ttctggggca	gaggcggagg	cgacgggctc	23280
tgcgggttcg	gcgggcggct	ggcagagccc	cttcgcggtt	cgggggtgcg	ctcctggcgg	23340
cgctgctctg	actgacttcc	tccgcggccg	gccattgtgt	tctcctaggg	agcaacaaca	23400
agcatggaga	ctcagccatc	gtcgccaaca	tgcctatctg	ccccgcgcgc	caccgccgac	23460
gagaaccagc	agcagaatga	aagcttaacc	gccccgcgcgc	ccagccccac	ctccgacgcc	23520
gcggccccag	acatgcaaga	gatggaggaa	tccatcgaga	ttgacctggg	ctacgtgacg	23580
cccgcggagc	acgaggagga	gctggcagcg	cgcttttcag	ccccggaaga	gaaccaccaa	23640
gagcagccag	agcaggaagc	agagaacgag	cagaaccagg	ctgggcacga	gcattggcgac	23700
tacctgagcg	gggcagagga	cgtgctcatc	aagcatctgg	cccgcctaat	catcatcgtc	23760
aaggacgcgc	tgctcgaccg	cgcgcagggtg	cccctcagcg	tggcgagact	cagccgcgcc	23820
tacgagcgca	acctcttctc	gccgcgcgtg	ccccccaagc	gccagcccaa	cggcacctgt	23880
gagcccaacc	cgcgcctcaa	cttctacccg	gtcttcgcgg	tggccgaggc	cctggccacc	23940
taccacctct	ttttcaagaa	ccaaaggatc	cccgtctcct	gccgcgcgca	ccgcacccgc	24000
gccgacgccc	tgctcaacct	gggccccggc	gccgcctac	ctgatatcac	ctccttggaa	24060
gaggttccca	agatcttcga	gggtctgggc	agcgacgaga	ctcgggcccgc	gaacgctctg	24120
caaggaagcg	gagaggagca	tgagcaccac	agcgccctgg	tggagtggga	aggcgacaac	24180
gcgcgcctgg	cggctcctcaa	gcgcacggtc	gagctgaccc	acttcgccta	cccggcgctc	24240
aacctgcccc	ccaagggtcat	gagcgccgtc	atggaccagg	tgctcatcaa	gcgcgcctcg	24300

cccctctcgg agggaggagat gcaggacccc gagagtctcg acgaggggcaa gcccgtaggtc 24360
 agcgacgagc agctggcgcg ctggctggga gcgagtagca cccccagag cctggaagag 24420
 cggcgcaagc tcatgatggc cgtggctcctg gtgaccgtgg agctggagtg tctgcgccgc 24480
 ttctttgccg acgcgagac cctgcgcaag gtcgaggaga acctgacta cctcttcagg 24540
 cacgggttcg tgcgccaggc ctgcaagatc tccaacgtgg agctgaccaa cctgggtctcc 24600
 tacatgggca tcttcgacga gaaccgcctg gggcaaaacg tgctgcacac caccctgcgc 24660
 ggggaggccc gccgcgacta catccgcgac tgcgtctacc tgtacctctg ccacacctgg 24720
 cagacgggca tgggcgtgtg gcagcagtgc ctggaggagc agaacctgaa agagctctgc 24780
 aagctcctgc agaagaacct caaggccctg tggaccgggt tcgacgagcg taccaccgcc 24840
 tcggacctgg ccgacctcat cttccccgag cgcttgcggc tgacgctgcg caacgggctg 24900
 cccgacttta tgagccaaag catgttgcaa aactttcgtc ctttcatcct cgaacgctcc 24960
 gggatcctgc ccgccacctg ctccgcgctg cctcggact tcgtgccgct gaccttccgc 25020
 gagtgcctcc cgccgctctg gagccactgc tacttgctgc gcctggccaa ctacctggcc 25080
 taccactcgg acgtgatcga ggacgtcagc ggcgagggtc tgctggagtg ccactgccgc 25140
 tgcaacctct gcacgccga ccgctccctg gcctgcaacc ccagctgct gagcgagacc 25200
 cagatcatcg gcaccttcga gttgcaaggc cccggcgacg gcgagggcaa ggggggtctg 25260
 aaactcacc cggggctgtg gacctcggcc tacttgcgca agttcgtgcc cgaggactac 25320
 catcccttcg agatcagggt ctacgaggac caatcccagc cgccaaggc cgagctgtcg 25380
 gcctgcgtca tcaaccaggg ggccatcctg gcccaattgc aagccatcca gaaatcccgc 25440
 caagaatttc tgctgaaaaa gggccacggg gtctacttgg accccagac cggagaggag 25500
 ctcaacccca gcttccccca ggatgccccg aggaagcagc aagaagctga aagtggagct 25560
 gccgccgccg gaggatttgg aggaagactg ggagagcagt caggcagagg agggaggagat 25620
 ggaagactgg gacagcactc aggcagagga ggacagcctg caagacagtc tggaggagga 25680
 agacgaggtg gaggaggcag aggaagaagc agccgccgcc agaccgtcgt cctcggcgga 25740
 gaaagcaagc agcacggata ccatctccgc tccgggtcgg ggtcgcggcg gccggggcca 25800
 cagtaggtgg gacgagaccg ggcgttccc gaacccacc acccagaccg gtaagaagga 25860
 gcggcagggg tacaagtcct ggcgggggca caaaaacgcc atcgtctcct gcttgcaagc 25920
 ctgcgggggc aacatctcct tcaccggcg ctacctgctc ttccaccgcg ggggtgaactt 25980

cccccgcaac atctttgcatt actaccgtca cctccacagc ccctactact gtttccaaga 26040
agaggcagaa acccagcagc agcagaâaac cagcggcagc agcagctaga aaatccacag 26100
cgggcgagcagg tggactgagg atcgcggcga acgagccggc gcagaccccg gagctgagga 26160
accggatctt tcccacccctc tatgccatct tccagcagag tcgggggcag gagcaggaac 26220
tgaaagtcaa gaaccgttct ctgcgctcgc tcacccgcag ttgtctgtat cacaagagcg 26280
aagaccaact tcagcgcaact ctcgaggacg ccgaggctct cttcaacaag tactgcgcgc 26340
tcactcttaa agagtagccc gcgcccgcac acacacggaa aaaggcggga attaggtcac 26400
cacctgcgcc cttcgcccga ccatcatgag caaagagatt cccacgcctt acatgtggag 26460
ctaccagccc cagatggggc tggcggcccg cgccggccag gactactcca cccgcagaa 26520
ctggctcagt gccgggcccg cgatgatctc acgggtgaat gacatccgcg cccaccgaaa 26580
ccagatactc ctagaacagt cagcgatcac cgccacgccc cgccatcacc ttaatccgcg 26640
taattggccc gccgccctgg tgtaccagga aattccccag cccacgaccg tactacttcc 26700
gcgagacgcc caggccgaag tcagctgac taactcaggt gtccagctgg ccggcggcgc 26760
cgccctgtgt cgtcacgcc ccgctcaggg tataaagcgg ctgggtgatcc gaggcagagg 26820
cacacagctc aacgacgagg tggtgagctc ttcgctgggt ctgcgacctg acggagtctt 26880
ccaactcgcc ggatcgggga gatcttcctt cagcctcgt caggccgtcc tgactttgga 26940
gagttcgctc tcgcagcccc gctcgggcgg catcggcact ctccagttcg tggaggagtt 27000
cactccctcg gtctacttca accccttctc cggctcccc ggccactacc cggacgagtt 27060
catcccgaac ttcgacgcca tcagcgagtc ggtggacggc tacgattgaa tgtcccatgg 27120
tggcgcagct gacctagctc ggcttcgaca cctggaccac tgccgcccgt tccgctgctt 27180
cgctcgggat ctgcgcgagt ttgcctactt tgagctgccc gaggagcacc ctcaggggccc 27240
agcccacgga gtgcggatca tcgtcgaagg gggcctcgac tcccacctgc ttcggatctt 27300
cagccagcga ccgatcctgg tcgagcgcca acaaggacag acccttctta ctttgtactg 27360
catctgcaac caccocggcc tgcagtaaag tctttgttgt ctgctgtgta ctgagtataa 27420
taaaagctga gatcagcgac tactccggac tcgattgtgg tgttctgct atcaaccgg 27480
ccctgttctt caccgggaac gagaccgagc tcagctcca gtgtaagccc cacaagaagt 27540
acctcacctg gctgttccag ggctccccga tcgcccgtgt caaccactgc gacaacgacg 27600
gagtcctgct gagcggccct gccaacctta ctttttccac ccgcagaagc aagctccagc 27660
tcttccaacc cttcctcccc gggacctatc agtgcgtctc aggacctgc catcacacct 27720

tccacctgat	cccgaatacc	acagcgccgc	tccccgctac	taacaaccaa	actaccacc	27780
aacgccaccg	tgcgacctt	tcctctgaat	ctaataccac	taccggaggt	gagctccgag	27840
gtcgaccaac	ctctgggatt	tactacggcc	cctgggaggt	ggtgggggta	atagcgctag	27900
gcctagtgtc	gggtgggctt	ttggttctct	gctacctata	cctcccttgc	tggtcgtact	27960
tagtgggtgt	gtgttgctgg	tttaagaaat	ggggaagatc	accctagtga	gctgcggtgc	28020
gctgggtggc	gtgttgcttt	cgattgtggg	actgggcggc	gcggctgtag	tgaaggagaa	28080
ggccgatccc	tgcttgcat	tcaatcccaa	caaatgccag	ctgagttttc	agcccgatgg	28140
caatcgggtc	gcggtactga	tcaagtgcgg	atgggaatgc	gagaacgtga	gaatcgagta	28200
caataacaag	actcggaa	atactctcgc	gtccgtgtgg	cagcccgggg	accccgagt	28260
gtacaccgtc	tctgtccccg	gtgctgacgg	ctccccgcgc	accgtgaata	atactttcat	28320
ttttgcgcac	atgtgcaaca	cggtcatgtg	gatgagcaag	cagtacgata	tgtggccccc	28380
cacgaaggag	aacatcgtgg	tcttctccat	cgcttacagc	ctgtgcacgg	cgctaatac	28440
cgctatcgtg	tgctgagca	ttcacatgct	catcgctatt	cgccccagaa	ataatgccga	28500
gaaagagaaa	cagccataac	acgttttttc	acacaccttg	tttttacaga	caatgcgtct	28560
gttaaatttt	ttaaacttg	tgctcagtat	tgcttatgcc	tctggttatg	caaacatata	28620
gaaaaccctt	tatgtaggat	ctgatgggtac	actagagggg	acccaatcac	aagccaaggt	28680
tgcatggtat	ttttatagaa	ccaacactga	tccagttaaa	ctttgtaagg	gtgaattgcc	28740
gcgtacacat	aaaactccac	ttacatttag	ttgcagcaat	aataatctta	cacttttttc	28800
aattacaaaa	caatatactg	gtacttatta	cagtacaaac	tttcatacag	gacaagataa	28860
atattatact	gttaaggtag	aaaatcctac	cactcctaga	actaccacca	ccaccactac	28920
tgcaaagccc	actgtgaaaa	ctacaactag	gaccaccaca	actacagaaa	ccaccaccag	28980
cacaacactt	gctgcaacta	cacacacaca	cactaagcta	accttacaga	ccactaatga	29040
tttgatcgcc	ctgctgcaaa	aggggggataa	cagcaccact	tccaatgagg	agatacccaa	29100
atccatgatt	ggcattattg	ttgctgtagt	ggtgtgcatg	ttgatcatcg	ccttgtgcat	29160
ggtgtactat	gccttctgct	acagaaagca	cagactgaac	gacaagctgg	aacacttact	29220
aagtgttgaa	ttttaatttt	ttagaaccat	gaagatccta	ggccttttta	gtttttctat	29280
cattacctct	gctctttgtg	aatcagtgga	tagagatggt	actattacca	ctggttctaa	29340
ttatacactg	aaagggccac	cctcagggtat	gctttcgtgg	tattgctatt	ttggaactga	29400

cactgatcaa	actgaattat	gcaattttca	aaaaggcaaa	acctcaaact	ctaaaatctc	29460
taattatcaa	tgcaatggca	ctgatctgat	actactcaat	gtcacgaaag	catatggtgg	29520
cagttattat	tgccctggac	aaaacactga	agaaatgatt	ttttacaaag	tggaagtggg	29580
tgatcccact	acaccacca	ccaccacaac	tattcatacc	acacacacag	aacaaacacc	29640
agaggcaaca	gaagcagagt	tggccttcca	ggttcacgga	gattcctttg	ctgtcaatac	29700
ccctacaccc	gatcagcggg	gtccggggcc	gctagtcagc	ggcattgtcg	gtgtgctttc	29760
gggattagca	gtcataatca	tctgcatggt	catttttgct	tgtgtctata	gaaggettta	29820
ccgacaaaaa	tcagaccac	tgctgaacct	ctatgtttaa	ttttttccag	agccatgaag	29880
gcagttagcg	ctctagtttt	ttgttctttg	attggcattg	tttttaatat	taaaattacc	29940
agagttagct	ttattaaaca	tggttaatgta	actgaaggag	ataacatcac	actagcaggt	30000
gtagaagggtg	ctcaaacac	cacctggaca	aaataccatc	taggatggag	agatatgtgc	30060
acctggaatg	taacttatta	ttgcatagga	gttaatctta	ccattgttaa	cgctaacca	30120
tctcagaatg	gggttaattaa	aggacagagt	gttagtgtga	ccagtgtatg	gtactatacc	30180
cagcatagtt	ttaactacaa	cattactgtc	ataccactgc	ctacgcctag	cccacctagc	30240
actaccacac	agacaaccac	atacagtaca	tcaaatcagc	ctaccacac	tacagcagca	30300
gaggttgcca	gctcgtctgg	gggtccgagt	gcatttttga	tggtggcccc	atctagcagt	30360
cccactgcta	gtaccaatga	gcagactact	gaatttttgt	ccactgtcga	gagccacacc	30420
acagctacct	ccagtgcctt	ctctagcacc	gccaatctct	cctcgctttc	ctctacacca	30480
atcagccccg	ctactactcc	tagccccgct	cctcttccca	ctccccgaa	gcaaacagac	30540
ggcggcatgc	aatggcagat	cacctgtctc	attgtgatcg	gggttggtcat	cctggccgtg	30600
ttgctctact	acatcttctg	ccgcgcgatt	cccaacgcgc	accgcaagcc	ggcctacaag	30660
cccacgttta	tcgggcagcc	ggagccgctt	caggtggaag	ggggtctaag	gaatcttctc	30720
ttctctttta	cagtatggtg	attgaactat	gattcctaga	caattcttga	tcactattct	30780
tatctgcctc	ctccaagtct	gtgccaccct	cgctctgggtg	gccaacgcca	gtccagactg	30840
tattggggcc	ttcgctctct	acgtgctctt	tgcttctgct	acctgcattc	gctgctgtag	30900
catagtctgc	ctgcttatca	ccttcttcca	gttcattgac	tggatctttg	tgcgcatcgc	30960
ctacctgcgc	caccaccccc	agtaccgcga	ccagcgagtg	gcgcagctgc	tcaggctcct	31020
ctgataagca	tgcgggctct	gtactttctc	gcgcttctgc	tgtagtgct	ccccgctccc	31080
gtcgaccccc	ggccccccac	tcagtcccc	gaggagggtc	gcaaatgcaa	attccaagaa	31140

ccctggaaat	tcttcaaagt	ctaccgcca	aatcagaca	tgcattcccag	ctggatcatg	31200
atcattggga	tgtgaacat	tctggcctgc	accctcatct	cctttgtgat	ttaccctctgc	31260
tttgactttg	gttggaactc	gccagaggcg	ctctatctcc	cgcctgaacc	tgacacacca	31320
ccacagcagc	aacctcaggc	acaagcacta	ccaccaccac	agcctaggcc	acaatacatg	31380
cccatattag	actatgaggc	cgagccacag	cgacccatgc	tccccgtat	tagttacttc	31440
aatctaaccg	gcggagatga	ctgaccact	ggccaataac	aacgtcaacg	accttctcct	31500
ggacatggac	ggcgcgcct	cggagcagcg	actcgccaa	cttcgcattc	gtcagcagca	31560
ggagagagcc	gtcaaggagc	tgcaggacgg	catagccatc	caccagtgc	agagaggcat	31620
cttctgcctg	gtgaaacagg	ccaagatctc	ctacgaggtc	accagaccg	accatcgct	31680
ctcctacgag	ctcctgcagc	agcgccagaa	gttcacctgc	ctggctggag	tcaaccccat	31740
cgtcatcacc	cagcagtcgg	gcgataccaa	ggggtgcac	cactgctcct	gcgactcccc	31800
cgactgcgtc	cacactctga	tcaagaccct	ctgcggcctc	cgcgacctcc	tccccatgaa	31860
ctaatacccc	ccttatccag	tgaataaag	atcatattga	tgatgattta	aataaaaaaa	31920
ataatcattt	gatttgaaat	aaagatacaa	tcatattgat	gatttgagtt	taacaaaaat	31980
aaagaatcac	ttacttgaaa	tctgatacca	ggctctctgc	catgttttct	gccaacacca	32040
cctcactccc	ctcttcccag	ctctggtaact	gcaggccccg	gcgggctgca	aacttctctc	32100
acacgctgaa	ggggatgtca	aattcctcct	gtccctcaat	cttcatttta	tcttctatca	32160
gatgtccaaa	aagcgcgtcc	gggtggatga	tgacttcgac	cccgtctacc	cctacgatgc	32220
agacaacgca	cgcaccgtgc	ccttcacaa	cccccccttc	gtctcttcag	atggattcca	32280
agagaagccc	ctgggggtgt	tgccctgcg	actggctgac	cccgtcacca	ccaagaacgg	32340
ggaaatcacc	ctcaagctgg	gagaggggg	ggacctcgac	tcgtcgggaa	aactcatctc	32400
caacacggcc	accaaggccg	cgcgccctct	cagtatttca	aacaacacca	tttcccttaa	32460
aactgctgcc	cctttctaca	acaacaatgg	aactttaagc	ctcaatgtct	ccacaccatt	32520
agcagtattt	cccacattta	acactttagg	cataagtctt	ggaaacggtc	ttcagacttc	32580
aaataagttg	ttgactgtac	aactaactca	tctcttaca	ttcagctcaa	atagcatcac	32640
agtaaaaaca	gacaaagggc	tatatattaa	ctccagtggg	aacagaggac	ttgaggctaa	32700
tataagccta	aaaagaggac	tagtttttga	cggtaatgct	attgcaacat	atattggaaa	32760
tggttagac	tatggatctt	atgatagtga	tggaaaaaca	agaccggtaa	ttacaaaaat	32820

tggagcagga ttaaattttg atgctaacaa agcaatagct gtcaaactag gcacaggttt 32880
 aagttttgac tccgctggtg cöttgacágc tggaaacaaa caggatgaca agctaact 32940
 ttggactacc cctgacccaa gccctaattg tcaattactt tcagacagag atgccaaatt 33000
 tactctctgt cttacaaaat gcggtagtca aatactagge actgtggcag tggcggtgt 33060
 tactgtagga tcagcactaa atccaattaa tgacacagtc aaaagcgcca tagttttcct 33120
 tagatttgat tccgatggtg tactcátgtc aaactcatca atggtaggtg attactggaa 33180
 ctttagggag ggacagacca ctcaaagtgt agcctataca aatgctgtgg gattcatgcc 33240
 aaatataggt gcatatccaa aaacccaaag taaaacacct aaaaatagca tágtcagtca 33300
 ggtatattta actggagaaa ctactatgcc aatgacacta accataactt tcaatggcac 33360
 tgatgaaaaa gacacaaccc cagtttagcac ctactctatg actttttacat ggcagtggac 33420
 tggagactat aaggacaaaa atattacctt tgctaccaac tcattctctt tttcctacat 33480
 cgcccaggaa taatcccacc cagcaagcca accccttttc ccaccacctt tgtctatatg 33540
 gaaactctga aacagaaaaa taaagttcaa gtgttttatt gaatcaacag ttttacagga 33600
 ctcgagcagt tatttttcct ccaccctccc aggacatgga ataccacc ctctccccc 33660
 gcacagcctt gaacatctga atgccattgg tgatggacat gcttttggtc tccacgttcc 33720
 acacagtttc agagcgagcc agtctcggat cggtcagggg gatgaaaccc tccgggcact 33780
 cccgcacctg cacctcacag ctcaacagct gaggattgtc ctcggtggtc gggatcacgg 33840
 ttátctggaa gaagcagaag agcggcggtg ggaatcatag tccgcgaacg ggatcggccg 33900
 gtggtgtcgc atcaggcccc gcagcagtcg ctgccccgc cgctccgtca agctgctgct 33960
 caggggggtc ggggtccaggg actccctcag catgatgcc acggccctca gcátcagtcg 34020
 tctggtgcgg cgggcgcagc agcgcatgcg aatctcgctc aggtcactgc agtacgtgca 34080
 acacaggacc accaggttgt tcaacagtcc atagttcaac acgctccagc cgaaactcat 34140
 cgcggggaagg atgctaccca cgtggccgtc gtaccagatc ctcaggtaaa tcaagtggcg 34200
 ctccctccag aagacgctgc ccattgtacat gatctccttg ggcattgtgg ggttcaccac 34260
 ctcccggtac cacatcccc tctgggtgaa catgcagccc cggatgatcc tgcggaacca 34320
 cagggccagc accgccccgc ccgcatgca gcgaagagac cccggatccc ggcaatgaca 34380
 atggaggacc caccgctcgt acccgtggat catctgggag ctgaacaagt ctatgttggc 34440
 acagcacagg catatgctca tgcátctctt cagcactctc agtcctcgg gggtcaaaaac 34500
 catatcccag ggcacgggga actcttgag gacagcgaac cccgcagaac agggcaatcc 34560

tcgcacataa cttacattgt gcatggacag ggtatcgcaa tcaggcagca ccgggtgatc 34620
 ctccaccaga gaagcgcggg tctcgggtctc ctacacagcgt ggtaaggggg ccggccgata 34680
 cgggtgatgg cgggacgcgg ctgatcgtgt tctcgaccgt gtcacatgac agttgtcttc 34740
 ggacatttct gtacttgctg tagcagaacc tgggtccgggc gctgcacacc gatcgccggc 34800
 ggcggtctcg gcgcttgga cgcctcgggtgt taaagttgta aaacagccac tctctcagac 34860
 cgtgcagcag atctagggcc tcaggagtga tgaagatccc atcatgcctg atagctctga 34920
 tcacatcgac caccgtggaa tgggccaggc ccagccagat gatgcaattt tgttgggttt 34980
 cgggtgacggc gggggaggga agaacaggaa gaaccatgat taacttttaa tccaaacggc 35040
 ctcgagcac ttcaaaatga aggtcacgga gatggcacct ctgcccccg ctgtgttggc 35100
 ggaaaataac agccagggtca aagggtgatac ggttctcgag atgttccacg gtggcttcca 35160
 gcaaagcctc cacgcgcaca tccagaaaca agacaatagc gaaagcggga ggggtctcta 35220
 attcctcaac catcatgtta cactctgca ccatcccg ataattttca ttttccagc 35280
 cttgaatgat tcgaactagt tcttgagga aatccaagc agccatgata aaaagctcgc 35340
 gcagagcacc ctccaccggc attcttaagc acacctcat aattccaaga tattctgctc 35400
 ctggttcacc tgcagcagat tgacaagcgg aatatcaaaa tctctgccgc gatccctgag 35460
 ctctccctc agcaataact gtaagtactc tttcatatcg tctccgaaat ttttagccat 35520
 aggaccccca ggaataagag aagggaagc cacattacag ataaaccgaa gtcccccca 35580
 gtgagcattg ccaaatgtaa gattgaaata agcatgctgg ctagaccgg tgatatcttc 35640
 cagataactg gacagaaaat cgggtaagca atttttaaga aaatcaacaa aagaaaaatc 35700
 ttccagggtgc acgtttaggc cctcgggaac aacgatggag taagtgaag ggggtgcgtc 35760
 cagcatggtt agttagctga tctgtaaaaa aacaaaaaat aaaacattaa accatgctag 35820
 cctggcgaac aggtgggtaa atcgttctct ccagcaccag gcaggccacg ggggtctccg 35880
 cgcgaccctc gtaaaaattg tcgctatgat tgaaaaccat cacagagaga cgttcccggt 35940
 ggccggcgtg aatgattcga gaagaagcat acaccccg aacattggag tccgtgagt 36000
 aaaaaagcg gccgaggaag caatgaggca ctacaacgct cactctcaag tccagcaaag 36060
 cgatgccatg cggatgaagc acaaaatttt cagggtgcgtg aaaaatgtaa ttactccct 36120
 cctgcacagg cagcgaagct cccgatecct ccagatacac atacaaagcc tcagcgtcca 36180
 tagcttaccg agcggcagca gcagcggcac acaacaggcg caagagtcag agaaaagact 36240

gagctctaac ctgtccgccc gctctctgct caatatatag cccagatct acactgacgt 36300
 aaaggccaaa gtctaaaaat acccgccaaa taatcacaca cgccagcac acgcccagaa 36360
 accggtgaca cactcagaaa aatacgcgca ctctctcaaa cggccaaact gccgtcattt 36420
 ccgggttccc acgtacgtc atcaaaacac gactttcaaa ttccgtcgac cgtaaaaaac 36480
 atcacccgcc ccgccctaa cggtcgcccgc tcccgagcc aatcaccttc ctccctcccc 36540
 aaattcaaac agctcatttg catattaacg cgcacaaaaa gtttgaggta tattattgat 36600
 gatg 36604

<210> 3
 <211> 36535
 <212> DNA
 <213> chimpanzee adenovirus serotype Pan7

<400> 3
 catcatcaat aatatacctc aaacttttgg tgcgcgttaa tatgcaaatg agctgtttga 60
 atttggggag ggaggaaggt gattggcga gagacgggag accgttaggg gcggggcggg 120
 tgacgttttt aatacgtggc cgtgaggcgg agccggtttg caagttctcg tgggaaaagt 180
 gacgtcaaac gaggtgtggt ttgaacacgg aaatactcaa ttttcccgcg ctctctgaca 240
 ggaaatgagg tgttttctggg cggatgcaag tgaaaacggg ccattttcgc gcgaaaactg 300
 aatgaggaag tgaaaatctg agtaatttcg cgtttatggc agggaggagt atttgccgag 360
 ggccgagtag actttgaccg attacgtggg ggtttcgatt accgtatttt tcacctaaat 420
 ttccgcgtac ggtgtcaaag tccggtgttt ttacgtaggc gtcagctgat cggcagggta 480
 tttaaacctg cgctctctag tcaagaggcc actcttgagt gccagcgagt agagttttct 540
 cctccgcgcc gcgagtcaga tctacacttt gaaagatgag gcacctgaga gacctgccc 600
 gtaatgtttt cctggctact gggaacgaga ttctggaatt ggtggtggac gccatgatgg 660
 gtggcgacc ccttgagccc cctaccccat ttgaggcgcc ttcgctgtac gatttgatg 720
 atctggaggt ggatgtgccc gagaacgacc ccaacgagga ggcggtgaat gatttgttta 780
 gcgatgccgc gctgctggct gccgagcagg ctaatacggc ctctggctca gacagcgatt 840
 cctctctcca taccctgaga cccggcagag gtgagaaaaa gatccccgag cttaaagggg 900
 aagagctcga cctgcgctgc tatgaggaat gcttgccctc gagcgatgat gaggaggacg 960
 aggaggcgat tcgagctgca tcgaaccagg gagtgaaagc tgcgggagaa agctttagcc 1020
 tggactgtcc tactctgccc ggacacggct gtaagtcttg tgaatttcat cgcataaata 1080

ctggagataa gaatgtgatg tgtgccctgt gctatatgag agcttacaac cattgtgttt	1140
acagtaagtg tgattaactt tagttgggaa ggcagagggg gactgggtgc tgactggttt	1200
atztatgtat atgttttttt atgtgtagggt cccgtctctg acgtagatga gacccccact	1260
tcagagtgc tttcatcacc ccagaaatt ggcgaggaac cgccgaaga tattattcat	1320
agaccagttg cagtgtgagt caccgggagg agagcagctg tggagagttt ggatgacttg	1380
ctacaggggtg gggatgaacc tttggacttg tgtaccggga aacgcccag gcactaagtg	1440
ccacacatgt gtgtttactt aaggtgatgt cagtatttat aggggtgtgga gtgcaataaa	1500
atccgtgttg actttaagtg cgtggtttat gactcagggg tggggactgt gggatatataa	1560
gcaggtgcag acctgtgtgg tcagttcaga gcaggactca tggagatctg gacgggtcttg	1620
gaagactttc accagactag acagctgcta gagaactcat cggagggggg ctcttacctg	1680
tggagattct gcttcgggtg gcctctagct aagctagtct atagggccaa acaggattat	1740
aaggatcaat ttgaggatat tttgagagag tgtcctggta tttttgactc tctcaacttg	1800
ggccatcagt ctacttttaa ccagagtatt ctgagagccc ttgacttttc tactcctggc	1860
agaactaccg ccgcggtagc cttttttgcc tttatccttg acaaaggag tcaagaaacc	1920
catttcagca gggattaccg tctggactgc ttagcagtag ctttgtggag aacatggagg	1980
tgccagcgcc tgaatgcaat ctccggctac ttgccagtae agccggtaga cacgctgagg	2040
atcctgagtc tccagtcacc ccaggaacac caacgccgc agcagccgca gcaggagcag	2100
cagcaagagg aggaggagga tcgagaagag aaccgagag ccgggtctgga cctccgggtg	2160
gcggaggagg aggagtagct gacttgtttc ccgagctgcg ccgggtgctg actaggtctt	2220
ccagtggacg ggagaggggg attaagcggg agaggcatga ggagactagc cacagaactg	2280
aactgactgt cagtctgatg agccgcaggc gcccagaatc ggtgtggtgg catgaggttc	2340
agtcgcaggg gatagatgag gtctcgggtga tgcagagaa atattccctg gaacaagtca	2400
agacttgttg gttggagcct gaggatgatt gggaggtagc catcaggaat tatgccaaagc	2460
tggctctgaa gccagacaag aagtacaaga ttaccaaact gattaatatc agaaattcct	2520
gctacatttc agggaatggg gccgagggtg agatcagtac ccaggagagg gtggccttca	2580
gatgttgat gatgaatat taccggggg tgggtggcat ggaggagtc acctttatga	2640
acgcgaggt caggggtgat gggataatg ggggtggtctt tatggccaac accaagctga	2700
cagtgcacgg atgctccttc tttgggttca ataacatgt catcgaggcc tggggcagtg	2760
tttcagtgag gggatgcagc ttttcagcca actggatggg ggtcgtgggc agaaccaaga	2820

gcaaggtgtc agtgaagaaa tgcctgttcg agaggtgcc a cctgggggtg atgagcgagg	2880
gcgaagccaa agtcaaacac tgcgctcta ctgagacggg ctgctttgtg ctgatcaagg	2940
gcaatgccca agtcaagcat aacatgatct gtggggcctc ggatgagcgc ggctaccaga	3000
tgctgacctg cgccggtggg aacagccata tgctggccac cgtgcatgtg acctcgacc	3060
cccgaagac atggcccag ttcgagcaca acgtcatgac ccgatgcaat gtgcacctgg	3120
ggccccgcg aggcattgtc atgccctacc agtgcaacat gcaatttgtg aagggtgctgc	3180
tggagcccg tgcctgtcc agagtggcc tgacgggggt gtttgacatg aatgtggagc	3240
tgtggaaaat tctgagatat gatgaatcca agaccagggt ccgggcctgc gaatgaggag	3300
gcaagcacgc caggcttcag ccggtgtgtg tggagggtgac ggaggacctg cgacccgatc	3360
atttgggtgt gtccctgaac gggacggagt tcggctccag cggggaagaa tctgactaga	3420
gtgagtagtg tttgggggag gtggagggtc tgtatgaggg gcagaatgac taaaatctgt	3480
gtttttctgt gtgttcagc agcatgagcg gaagcgctc ctttgaggga ggggtattca	3540
gcccttatct gacggggcgt cccccctct gggcgggagt gcgtcagaat gtgatgggat	3600
ccacggtgga cggccggccc gtgcagccc cgaactcttc aacctgacc tacgcgacc	3660
tgagctcttc gtccgtggac gcagctgcc ccgcagctgc tgcttccgcc gccagcgccg	3720
tgcgcggaat ggccctgggc gccggctact acagctctct ggtggccaac tcgacttcca	3780
ccaataatcc cgccagcctg aacgaggaga agctgctgct gctgatggcc cagctcgagg	3840
ccctgaccca gcgcctgggc gagctgacct agcagggtggc tcagctgcag ggggagacgc	3900
gggcgcggtg tgccacggtg aaaaccaa ataaaaatgaa tcaataaata aacggagacg	3960
gttggtgatt ttaacacaga gtcttgaatc tttatttgat ttttcgcgcg cggtagggcc	4020
tggaccaccg gtctcgatca ttgagcacc ggtggatttt ttccaggacc cggtagagg	4080
gggcttgat gttaggtac atgggcatga gcccgctccg ggggtggagg tagctccatt	4140
gcagggcctc gtgctcgggg gtggtgttgt aaatcaccca gtcatagcag gggcgagg	4200
cgtggtgctg cagcatgtcc ttgaggagga gactgatggc cacgggcagc ccttggtgt	4260
agggtgtgac gaacctgttg agctgggagg gatgcatgcg gggggagatg agatgcatct	4320
tggcctggat cttgagattg gcgatgttcc cgccagatc ccgccccggg ttcattgtgt	4380
gcaggaccac cagcacggtg tatccggtgc acttggggaa tttgtcatgc aacttggag	4440
ggaaggcgtg aaagaatttg gagacgcctc tgtgaccgcc caggttttcc atgcactcat	4500

ccatgatgat	ggcgatgggc	cogtgggcgg	cggcctgggc	aaagacgttt	cgggggtcgg	4560
acacatcgta	gttgtgggtcc	tgggtgagct	cgatcataggc	catttttaatg	aatttggggc	4620
ggaggggtgcc	cgactggggg	acgaagggtgc	cctcgatccc	gggggcgtag	ttgccctcgc	4680
agatctgcat	ctcccaggcc	ttgagctcgg	agggggggat	catgtccacc	tgcggggcga	4740
tgaaaaaaac	ggtttccggg	gcgggggaga	tgagctgggc	cgaaagcagg	ttccggagca	4800
gctgggactt	gccgcagccg	gtggggccgt	agatgacccc	gatgaccggc	tgcaggtggt	4860
agttgagggg	gagacagctg	ccgtcctcgc	ggaggagggg	ggccacctcg	ttcatcatct	4920
cgcgcacatg	catgttctcg	cgcacgagtt	ccgccaggag	gcgctcgccc	cccagcgaga	4980
ggagctcttg	cagcgaggcg	aagtttttca	gcggcttgag	yccgtcggcc	atgggcattt	5040
tggagagggg	ctgttgcaag	agttccagac	ggtcccagag	ctcgggtgatg	tgctctaggg	5100
catctcgatc	cagcagacct	cctcgtttcg	cgggttgggg	cgactgcggg	agtagggcac	5160
caggcgatgg	gcgtccagcg	aggccagggt	ccggtccttc	cagggtcgca	gggtccgcgt	5220
cagcgtggtc	tccgtcacgg	tgaaggggtg	cgcgccgggc	tgggcgcttg	cgagggtgcg	5280
cttcaggctc	atccggctgg	tcgagaaccg	ctcccggtcg	gcgccctcgc	cgtcggccag	5340
gtagcaattg	agcatgagtt	cgtagttgag	cgcctcggcc	gcgtggccct	tggcgcgag	5400
cttacctttg	gaagtgtgtc	cgcagacggg	acagaggagg	gacttgaggg	cgtagagctt	5460
gggggcgagg	aagacggact	cgggggcgta	ggcgtccgcg	ccgcagctgg	cgcagacggt	5520
ctcgcactcc	acgagccagg	tgagggtcggg	ccgggttgggg	tcaaaaacga	ggtttcctcc	5580
gtgctttttg	atgcgtttct	tacctctggt	ctccatgagc	tcgtgtcccc	gctgggtgac	5640
aaagaggctg	tccgtgtccc	cgtagaccga	ctttatgggc	cggtcctcga	gcgggggtgcc	5700
gcggtcctcg	tcgtagagga	accccgccca	ctccgagacg	aaggcccggg	tccaggccag	5760
cacgaaggag	gccacgtggg	aggggtagcg	gtcgttgtec	accagcgggt	ccaccttctc	5820
cagggtatgc	aagcacatgt	ccccctcgtc	cacatccagg	aagggtgattg	gcttgtaagt	5880
gtaggccacg	tgaccggggg	tcccggccgg	gggggtataa	aagggggcgg	gcccctgctc	5940
gtcctcactg	tcttcgggat	cgctgtccag	gagcgccagc	tgttggggta	ggtattccct	6000
ctcgaaggct	ggcataacct	cggcactcag	gttgtcagtt	tctagaaacg	aggaggattt	6060
gatattgacg	gtgccgttgg	agacgccttt	catgagcccc	tcgtccatct	ggtcagaaaa	6120
gacgatcttt	ttgttgtcga	gcttggtggc	gaaggagccg	tagagggcgt	tggagaggag	6180
cttggcgatg	gagcgcatgg	tctgggtctt	ttccttgctg	gcgcgctcct	tggcgggcat	6240

gttgagctgc acgtactcgc gcgccacgca cttccattcg gggaaagacgg tggtagctc	6300
gtcgggcacg attctgaccc gccagccgcg gttgtgcagg gtgatgaggt ccacgctggt	6360
ggccacctcg ccgcgcaggg gctcgttggg ccagcagagg cgcgcgccct tgcgcgagca	6420
gaaggggggc agcgggtcca gcatgagctc gtcggggggg tcggcgcca cggatgaagat	6480
gccgggcaga agctcggggg cgaagttagct gatgcagggtg tccagatcgt ccagcgccgc	6540
ttgccagtcg cgcacggcca gcgcgcgctc gtaggggctg aggggcgtgc cccagggcat	6600
ggggtgcgtg agcgcggagg cgtacatgcc gcagatgtcg tagacgtaga ggggtcctc	6660
gaggacgccg atgtagggtg ggtagcagcg ccccccgcgg atgctggcgc gcacgtagtc	6720
gtacagctcg tgcgagggcg cgaggagccc cgtgccgagg ttggagcgtt gcggcttttc	6780
ggcgcggtag acgatctggc ggaagatggc gtgggagttg gaggagatgg tgggcctctg	6840
gaagatgttg aagtgggcgt ggggcaggcc gaccgagtc ctgatgaagt gggcgtagga	6900
gtcctgcagc ttggcgacga gctcggcggt gacgaggacg tccagggcgc agtagtcgag	6960
ggtctcttgg atgatgtcgt acttgagctg gcccttctgc ttccacagct cgcggttgag	7020
aaggaactct tcgcggctct tccagtactc ttcgaggggg aaccgcctct gatcggcacg	7080
gtaagagccc accatgtaga actgggtgac ggcttgtag gcgcagcagc ccttctccac	7140
ggggagggcg taagcttggt cggccttgcg caggagagtg tgggtgaggg cgaagggtgc	7200
gcgcaccatg accttgagga actggtgctt gaagtcgagg tcgtcgcagc cgccctgctc	7260
ccagagctgg aagtcctgct gcttcttgta ggccggggtg ggcaaagcga aagtaacatc	7320
gttgaagagg atcttgcccg cgcggggcat gaagttgcga gtgatgcgga aaggctgggg	7380
cacctcgcc cggttggtga tgacctgggc ggcgaggacg atctcgtcga agccgttgat	7440
gttgtgcccg acgatgtaga gttccacgaa tcgcgggcgg cccttaacgt ggggcagctt	7500
cttgagctcg tcgtagggtga gtcgggcggg gtcgctgagc ccgtgctgct cgaggggcca	7560
gtcggcgacg tgggggttgg cgctgaggaa ggaagtccag agatccacgg ccagggcggg	7620
ctgcaagcgg tcccgggtact gacggaactg ctggcccacg gccatttttt cgggggtgac	7680
gcagtagaag gtgcgggggt cgcctgcca gcggtccac ttgagctgga gggcgaggtc	7740
gtgggcgagc tcgacgagcg gcgggtcccc ggagagtttc atgaccagca tgaaggggac	7800
gagctgcttg ccgaaggacc ccatccaggt gtaggtttcc acatcgtagg tgaggaagag	7860
cctttcggtg cgaggatgcg agccgatggg gaagaactgg atctcctgcc accagttgga	7920

ggaatggctg ttgatgtgat ggaagtagaa atgccgacgg cgcgccgagc actcgtgctt	7980
gtgtttatac aagcgtccgc agtgctcgca acgctgcacg ggatgcacgt gctgcacgag	8040
ctgtacctgg gttcctttga cgaggaattt cagtgggcag tggagcgctg gcggtgcat	8100
ctggtgctgt actacgtcct ggccatcggc gtggccatcg tctgcctcga tgggtggcat	8160
gctgacgagc cgcgcgggga ggcaggcca gacttcggct cggacgggtc ggagagcgag	8220
gacgagggcg cgcaggccgg agctgtccag ggtcctgaga cgctgcggag tcaggtcagt	8280
gggcagcggc ggcgcgcggg tgaactgcag gagcttttcc agggcgcggc ggaggccaag	8340
atggtacttg atctccacgg cgcggttggg ggcgacgtcc acggcttgca gggccccgtg	8400
cccctggggc gccaccaccg tgccccgttt cttcttgggc gctgcttcca tgccggtcag	8460
aagcggcggc gaggacgcgc gccgggcggc aggggcggct cgggacccgg aggcaggggc	8520
ggcaggggca cgtcggcgcc gcgcgcgggc aggttctggg actgcgcccg gagaagactg	8580
gcgtgagcga cgacgcgacg gttgacgtcc tggatctgac gcctctgggt gaaggccacg	8640
ggacccgtga gtttgaacct gaaagagagt tcgacagaat caatctcggt atcgttgacg	8700
gcggcctgcc gcaggatctc ttgcacgtcg cccgagttgt cctggtaggc gatctcggtc	8760
atgaactgct cgatctctc ctctgaagg tctccgcggc cggcgcgctc gacggtggcc	8820
gcgaggtcgt tggagatgcy gcccatgagc tgcgagaagg cgttcatgcc ggcctcgttc	8880
cagacgcggc tgtagaccac ggctccgtcg gggtcgcgcg cgcgcacgac cacctgggcy	8940
aggttgagct cgacgtggcg cgtgaagacc gcgtagttgc agaggcgctg gtagaggtag	9000
ttgagcgtgg tggcgatgtg ctcggtgacg aagaagtaca tgatccagcy gcgagcggc	9060
atctcgctga cgtcgcccag ggcttccaag cgctccatgg cctcgtagaa gtccacggcy	9120
aagttgaaaa actgggagtt gcgcgcgag acggtcaact cctcctccag aagacggatg	9180
agctcagcga tgggtggcgcg cacctcgcgc tcgaaggccc cggggggctc ctcttcttcc	9240
atctcttct cctccactaa catctcttct acttctctct caggaggcgg cggcggggga	9300
ggggccctgc gtcgcggcg gcgcacgggc agacggtcga tgaagcgtc gatggtctcc	9360
ccgcgcggc gacgcacgtg ctcggtgacg gcgcgccgt cctcgcgggg ccgcagcgtg	9420
aagacgcgc cgcgcacctc cagggtggcg ccgggggggt ctccgttggg cagggaagag	9480
gcgctgacga tgcatcttat caattggccc gtagggactc cgcgcaagga cctgagcgtc	9540
tcgagatcca cgggatccga aaaccgctga acgaaggctt cgagccagtc gcagtcgcaa	9600
ggtaggctga gcccggtttc ttgttcttcg gggatttcgg gaggcggcg ggcgatgctg	9660

ctgggtgatga agttgaagta ggcggtcctg agacggcgga tggtaggcgag gagcaccagg	9720
tccttggggcc cggcttgctg gatgcgcaga cggtcggcca tgccccaggc gtggtcctga	9780
cacctggcga ggtccttgta gtagtcctgc atgagccgct ccacgggcac ctccctcctcg	9840
cccgcgcggc cgtgcatgcg cgtgagcccc aaccgcgct ggggctggac gagcgccagg	9900
tcggcgacga cgcgctcggc gaggatggcc tgctgtatct gggtaggggt ggtctggaag	9960
tcgtcgaagt cgacgaagcg gtggtaggct ccggtgttga tggatatagga gcagttggcc	10020
atgacggacc agttgacggt ctggtaggccc ggtcgacga gctcgtggtta cttgaggcgc	10080
gagtaggcgc gcgtgtcga gatgtagtcg ttgcagggtgc gcacgaggta ctggatatccg	10140
acgaggaagt gcggcgcgcg ctggcggtag agcggccatc gctcgtggc gggggcgccg	10200
ggcgcgaggc cctcgagcat gaggcggtag tagccgtaga tgtacctgga catccagggtg	10260
atgccggcg cggtggtgga ggcgcgcggg aactcgcgga cgcggttcca gatgttgcgc	10320
agcggcagga agtagttcat ggtggcgcgc gtctggcccc tgaggcgcg gcagtcgtgg	10380
atgctctaga catacgggca aaaacgaaag cggtcagcgc ctcgactccg tggcctggag	10440
gctaagcgaa cgggttgggc tgcgcgtgta ccccggttcg aatctcgaat caggctggag	10500
ccgcagctaa cgtggtactg gcaactccgt ctcgacccaa gcctgctaac gaaacctcca	10560
ggatacggag gcgggtcggt ttttggcctt ggtcgctggt catgaaaaac tagtaagcgc	10620
ggaaagcgac cggcccgat ggctcgctgc cgtagtcctg agaaagaatc gccagggttg	10680
cgttgcggtg tgccccggtt cgagcctcag cgctcgcggc cggccggatt ccgcggctaa	10740
cgtgggcgtg gctgccccgt cgtttccaag accccttagc cagccgactt ctccagttac	10800
ggagcgagcc cctctttttt ttgtgttttt gccagatgca tcccgtactg cggcagatgc	10860
gccccaccc tccacctcaa cgcgccctac cgccgcagca gcagcaacag ccggcgcttc	10920
tgcccccgcc ccagcagcag ccagccacta ccgcggcggc gcgcgtgagc ggagccggcg	10980
ttcagtatga cctggccttg gaagagggcg aggggctggc gcggctgggg gcgtcgtcgc	11040
cggagcggca cccgcgcgtg cagatgaaaa gggacgtcg cgaggcctac gtgcccgaagc	11100
agaacctgtt cagagacagg agcggcgagg agcccagga gatgcgcgcc tcccgtttcc	11160
acgcggggcg ggagctgcgg cgcggcctgg accgaaagcg ggtgctgagg gacgaggatt	11220
tcgaggcgga cgagctgacg gggatcagcc ccgcgcgcgc gcacgtggcc gcggccaacc	11280
tggtcacggc gtacgagcag accgtgaagg aggagagcaa cttccaaaaa tccttcaaca	11340

accacgtgcg cacgctgate gcgcgcgagg aggtgaccct gggcctgatg cacctgtggg 11400
 acctgctgga ggccatcgtg cagaacccca cgagcaagcc gctgacggcg cagctgttcc 11460
 tgggtgtgca gcacagtcgg gacaacgaga cggtcagggg ggcgctgctg aatatcaccg 11520
 agcccgaggg cgctggctc ctggacctgg tgaacattct gcagagcatc gtggtgcagg 11580
 agcgcgggct gccgctgtcc gagaagctgg cggctatcaa cttctcgggtg ctgagcctgg 11640
 gcaagtacta cgctaggaag atctacaaga ccccgtagct gcccatagac aaggagggtga 11700
 agatcgacgg gttttacatg cgcgatgaccc tgaagtgtct gaccctgagc gacgatctgg 11760
 ggggtgtaccg caacgacagg atgcaccgcg cgggtgagcgc cagccgccgg cgcgagctga 11820
 gcgaccagga gctgatgcac agcctgcagc gggccctgac cggggccggg accgaggggg 11880
 agagctactt tgacatgggc gcggacctgc gctggcagcc cagccgccgg gccttggaag 11940
 ctgccggcgg ttccccctac gtggaggagg tggacgatga ggaggaggag ggcgagtacc 12000
 tgggaagactg atggcgcgac cgtatttttg ctagatgcag caacagccac cgcctcctga 12060
 tcccgcgatg cgggcggcgc tgcagagcca gccgtccggc attaactcct cggacgattg 12120
 gaccagggcc atgcaacgca tcatggcgct gacgaccgc aatcccgagg cctttagaca 12180
 gcagcctcag gccaacggc tctcgcccat cctggaggcc gtggtgcctt cgcgctcgaa 12240
 cccacgcac gagaagggtc tggccatcgt gaacgcgctg gtggagaaca aggccatccg 12300
 cggcgacgag gccgggctgg tgtacaacgc gctgctggag cgcgtggccc gctacaacag 12360
 caccaacgtg cagacgaacc tggaccgcat ggtgaccgac gtgcgcgagg cgggtgtcgca 12420
 gcgcgagcgg ttccaccgcg agtcgaacct gggctccatg gtggcgctga acgccttcc 12480
 gagcacgcag cccgccaacg tgccccgggg ccaggaggac tacaccaact tcatcagcgc 12540
 gctgcggctg atgggtggcg aggtgcccc gagcgagggtg taccagtcgg ggcgggacta 12600
 cttcttccag accagtgcgc agggcttgca gaccgtgaac ctgagccagg ctttcaagaa 12660
 cttgcaggga ctgtggggcg tgcaggcccc ggtcggggac cgcgcgacgg tgtcagacct 12720
 gctgacgccg aactcgcgcc tgctgctgct gctggtggcg cccttcacgg acagcggcag 12780
 cgtgagccgc gactcgtacc tgggctacct gcttaacctg taccgcgagg ccatcgggca 12840
 ggcgcacgtg gacgagcaga cctaccagga gatcaccac gtgagccgcg cgctgggcca 12900
 ggaggacccg ggcaacctgg aggccaccct gaacttctg ctgaccaacc ggtcgcagaa 12960
 gatcccgccc cagtacgcgc tgagcaccga ggaggagcgc atcctgcgct acgtgcagca 13020
 gagcgtgggg ctgttctga tgcaggaggg ggccacgccc agcgcgcgc tcgacatgac 13080

cgcgcgcaac atggagccca gcatgtacgc tcgcaaccgc ccgttcatca ataagctgat	13140
ggactacttg catcgggcgg ccgccatgaa ctcggactac tttaccaacg ccatcttgaa	13200
cccgcaactgg ctcccgccgc ccgggttcta cacgggcgag tacgacatgc ccgaccccaa	13260
cgacgggttc ctgtgggacg acgtggacag cagcgtgttc tcgccgcgcc ccgccaccac	13320
cgtgtggaag aaagagggcg gggaccggcg gccgtcctcg gcgtgtccg gtcgcgcggg	13380
tgctgccgcg gcggtgcctg aggccgccag ccccttcccg agcctgccct tttcgtgaa	13440
cagcgtgcgc agcagcgagc tgggtcggct gacgcggccg cgcctgctgg gcgaggagga	13500
gtacctgaac gactccttgt tgaggcccg gcgcgagaag aacttcccca ataacgggat	13560
agagagcctg gtggacaaga tgagccgctg gaagacgtac gcgcacgagc acagggacga	13620
gccccgagct agcagcagcg caggcacccg tagacgccag cgacacgaca ggcagcgggg	13680
tctggtgtgg gacgatgagg attecgccga cgacagcagc gtgttggact tgggtgggag	13740
tgggtggtgg aaccgcgttc ctcacttgcg ccccggtatc gggcgccctga tgtaagaatc	13800
tgaaaaata aaaaacggta ctcaccaagg ccatggcgac cagcgtgcgt tcttctctgt	13860
tgttttagt agtatgatga ggcgcgtgta cccggagggt cctcctccct cgtacgagag	13920
cgtgatgcag caggcgggtg cggcggcgat gcagccccg ctggaggcgc cttacgtgcc	13980
cccgcggtac ctggcgcccta cggagggggcg gaacagcatt cgttactcgg agctggcacc	14040
cttgtagat accaccgggt tgtacctggg ggacaacaag tcggcggaca tcgcctcgct	14100
gaactaccag aacgaccaca gcaacttcct gaccaccgtg gtgcagaaca acgatttcac	14160
ccccacggag gccagcacc agaccatcaa ctttgacgag cgctcgcggt gggcgggcca	14220
gctgaaaacc atcatgcaca ccaacatgcc caacgtgaac gagttcatgt acagcaacaa	14280
gttcaaggcg cgggtgatgg tctcgcgcaa gacccccaat ggggtcgcgg tggatgagaa	14340
ttatgatggt agtcaggacg agctgactta cgagtgggtg gagtttgagc tgcccagggg	14400
caacttctcg gtgacctga ccatcgatct gatgaacaac gccatcatcg acaactactt	14460
ggcgggtggg cgtcagaacg ggggtgctgga gagcgacatc ggcgtgaagt tcgacacgcg	14520
caacttccgg ctgggctggg acctcgtgac cgagctggtg atgccggggc tgtacaccaa	14580
cgaggccttc cccccgaca tcgtcctgct gcccggtgc ggcgtggact tcaccgagag	14640
ccgcctcagc aacctgctgg gcatccgcaa gcggcagccc ttccaggagg gcttcagat	14700
cctgtacgag gacctggagg ggggcaacat ccccgcgctc ttggatgtcg aagcctatga	14760

gaaaagcaag gaggaggccg ccgcagcggc gaccgcagcc gtggccaccg cctctaccga	14820
ggtagcggggc gataattttg ctagcgccgc ggcagtggcc gaggcggctg aaaccgaaag	14880
taagatagtc atccagcccg tggagaagga cagcaaggac aggagctaca acgtgctcgc	14940
ggacaagaaa aacaccgcct accgcagctg gtacctggcc tacaactacg gcgaacccga	15000
gaagggcgtg cgctcctgga cgctgctcac cacctcggac gtcacctgcy gcgtggagca	15060
agtctactgg tcgctgcccg acatgatgca agaccgggtc accttcgct ccacgcgtca	15120
agttagcaac taccgggtgg tgggcgcga gtcctgccc gtctactcca agagcttctt	15180
caacgagcag gccgtctact cgcagcagct gcgcgccttc acctcgctca cgcagctctt	15240
caaccgcttc cccgagaacc agatcctcgt ccgcccgcgc gcgccacca ttaccaaccg	15300
cagtgaatac gttcctgctc tcacagatca cgggaccctg ccgctgcgca gcagtatccg	15360
gggagtccag cgcgtgaccg tcaactgacgc cagacgccgc acctgcccct acgtctacaa	15420
ggccctgggc gtagtcgcgc cgcgcgtcct ctgcagccgc accttctaaa aaatgtccat	15480
tctcatctcg cccagtaata acaccggttg gggcctgcgc gcgccagca agatgtacgg	15540
aggcgtcgc caacgtcca cgcaacacc cgtgcgcgtg cgcgggcact tccgcgtcc	15600
ctggggcgcc ctcaagggcc gcgtgcgtc gcgcaccacc gtcgacgacg tgatcgacca	15660
ggtagtgccc gacgcgcga actacacgcc cgcgcgcgc cccgcctcca ccgtggacgc	15720
cgtcatcgac agcgtggtgg ccgatgcgc ccggtacgcc cgcgcgaaga gccggcggcg	15780
gcgcacgcc ccggcgccacc ggagcaccac cgcctatgcgc gcggcgcgag ccttgctgcg	15840
cagggccagg cgcacgggac gcagggccat gctcagggcg gccagacgcg cggcctccg	15900
cagcagcagc gccggcagga cccgcagacg cgcggccacg gcggcgcgcg cggccatcgc	15960
cagcatgtcc cgcgcgcggc gcggcaacgt gtactgggtg cgcgacgcg ccaccgggtg	16020
gcgcgtgccc gtgcgcaccc gccccctcg cacttgaaga tgctgacttc gcgatgttga	16080
tgtgtcccag cggcgaggag gatgtccaag cgcaaataca aggaagagat gctccaggtc	16140
atcgcgcctg agatctacgg ccccgcggtg aaggaggaaa gaaagccccg caaactgaag	16200
cgggtcaaaa aggacaaaaa ggaggaggaa gatgtggacg gactggtgga gtttgtgcgc	16260
gagttcgcgc cccggcggcg cgtgcagtgg cgcgggcgga aagtgaacc ggtgctgcg	16320
cccggcacca cggtaggtctt caccgccggc gagcgttccg gctccgcctc caagcgtcc	16380
tacgacgagg tgtacgggga cgaggacatc ctgcagcagg cggtcgagcg tctgggcgag	16440
tttgettacg gcaagcgcag ccgccccgcg cccttgaaag aggaggcggg gtccatcccc	16500

ctggaccacg gcaacccac gccgagcctg aagccggtga ccctgcagca ggtgctgccg 16560
 agcgcgggcg cgcgccggg cttcaagcgc gaggcgggcg aggatctgta cccgaccatg 16620
 cagctgatgg tgcccaagcg ccagaagctg gaggacgtgc tggagcacat gaaggaggac 16680
 cccgaggtgc agcccgaggt caagggtgcg cccatcaagc aggtggcccc gggcctgggc 16740
 gtgcagaccg tggacatcaa gatccccacg gagcccatgg aaacgcagac cgagcccgtg 16800
 aagcccagca ccagcaccat ggagggtgag acggatccct ggatgccggc gccggcctcc 16860
 accactcgcc gaagacgcaa gtacggcgcg gccagcctgc tgatgccccaa ctacgcgctg 16920
 catccttcca tcatccccac gccgggctac cgcggcacgc gcttctaccg cggctacacc 16980
 agcagccgcc gcaagaccac caccgcccgc cgccgtcgtc gcaccgcgcg cagcagcacc 17040
 gcgacttccg ccgcccgcct ggtgcggaga gtgtaccgca gcgggcgcga gcctctgacc 17100
 ctgccgcgcg cgcgctacca cccgagcatc gccatttaac tctgccgtcg cctcctactt 17160
 gcagatatgg cctcacatg ccgcctccgc gtccccatta cgggctaccg aggaagaaag 17220
 ccgcgcgta gaaggctgac ggggaacggg ctgcgtcgcc atcaccaccg gcggcgcgcg 17280
 gccatcagca agcgggtggg gggaggcttc ctgccgcgc tgatccccat catcgccgcg 17340
 gcgatcgggg cgatccccgg catagcttcc gtggcggtgc aggcctctca gcgccactga 17400
 gacacagctt ggaaaatttg taataaaaaa atggactgac gctcctggtc ctgtgatgtg 17460
 tgtttttaga tggaagacat caatttttcg tccctggcac cgcgacacgg cacgcggccg 17520
 tttatgggca cctggagcga catcggaac agccaactga acggggggcg cttcaattgg 17580
 agcagtctct ggagcgggct taagaatttc ggggtccacgc tcaaaccta tggcaacaag 17640
 gcgtggaaca gcagcacagg gcaggcgctg agggaaaagc tgaaagagca gaacttccag 17700
 cagaagggtg tcgatggcct ggcctcgggc atcaacgggg tggtagacct ggccaaccag 17760
 gccgtgcaga aacagatcaa cagccgcctg gacgcggtcc cggccgcggg gtccgtggag 17820
 atgccccagg tggaggagga gctgcctccc ctggacaagc gcggcgacaa gcgacgcgt 17880
 cccgacgcgg aggagacgt gctgacgcac acggacgagc cggccccgta cgaggaggcg 17940
 gtgaaactgg gtctgcccac cacgcggccc gtggcgcttc tggccaccgg ggtgctgaaa 18000
 cccagcagca gcagccagcc cgcgacctg gacttgcttc cgctgcttc cggccctcc 18060
 acagtggcta agcccctgcc gccgggtggc gtcgcgtcgc gcgcccccg aggcgcgcc 18120
 caggcgaact ggcagagcac tctgaacagc atcgtgggtc tgggagtga gagtgtgaag 18180

cgccgcccgt gctattaaaa gacactgtag cgcttaactt gcttgtctgt gtgtatatgt 18240
 atgtccgccg accagaagga ggaagaggcg cgtcgccgag ttgcaagatg gccaccccat 18300
 cgatgctgcc ccagtgggcg tacatgcaca tcgccggaca ggacgcttcg gagtacctga 18360
 gtccgggtct ggtgcagttc gcccgcgcca cagacaccta cttcagttctg gggacaagt 18420
 ttaggaaccc caccggtggcg cccacgcacg atgtgaccac cgaccgcagc cagcggctga 18480
 cgctgcgctt cgtgcccgtg gaccgcgagg acaacaccta ctctacaaa gtgcgctaca 18540
 cgctggccgt gggcgacaac cgcgtgctgg acatggccag cacctacttt gacatccgcg 18600
 gcgtgctgga tcggggggccc agcttcaaac cctactccgg caccgcctac aacagcctgg 18660
 ctcccaaggg agcgcccaac acttgccagt ggacatataa agctggtgat actgatacag 18720
 aaaaaaccta tacatatgga aatgcacctg tgcaaggcat tagcattaca aaggatggta 18780
 ttcaacttgg aactgacagc gatggtcagg caatctatgc agacgaaact tatcaaccag 18840
 agcctcaagt ggggatgctt gaatggcatg acatcactgg tactgatgaa aaatatggag 18900
 gcagagctct taagcctgac accaaaatga agccttgcta tggttctttt gccaaagcta 18960
 ccaataaaga aggaggccag gcaaatgtga aaaccgaaac aggcgggtacc aaagaatatg 19020
 acattgacat ggcattcttc gataatcgaa gtgcagctgc cgccggccta gcccagaaa 19080
 ttgttttgta tactgagaat gtggatctgg aaactccaga taccatatt gtatacaagg 19140
 caggtagaga tgacagtagc tcttctatca atttgggtca gcagtcctatg cccaacagac 19200
 ccaactacat tggcttcaga gacaacttta tcggtctgat gtactacaac agcactggca 19260
 atatgggtgt actggctgga caggcctccc agctgaatgc tgtgggtggac ttgcaggaca 19320
 gaaacaccga actgtctac cagctcttgc ttgactctct gggtgacaga accaggtatt 19380
 tcagtatgtg gaatcaggcg gtggacagtt atgacccga tgtgcgcatt attgaaaatc 19440
 acggtgtgga ggatgaactt cctaactatt gcttccccct ggatgctgtg ggtagaactg 19500
 atacttacca gggaaattaag gccaatggtg ataatcaaac cacctggacc aaagatgata 19560
 ctgttaatga tgctaataa ttgggcaagg gcaatccttt cgccatggag atcaacatcc 19620
 aggccaacct gtggcggaac ttctctacg cgaacgtggc gctgtacctg cccgactcct 19680
 acaagtacac gccggccaac atcacgtgc ccaccaaac caacacctac gattacatga 19740
 acggccgctg ggtggcgccc tcgctgggtg acgcctacat caacatcggg gcgcgctggg 19800
 cgctggaccc catggacaac gtcaaccct tcaaccacca ccgcaacgcg ggcctgcgat 19860
 accgctccat gctcctgggc aacgggcgct acgtgccctt ccacatccag gtgccccaaa 19920

agtttttgcg catcaagagc ctctgtctcc tgcccgggtc ctacacctac gagtggaaact 19980
 tccgcaagga cgtcaacatg atcctgcaga gctccctcgg caacgacctg cgcacggagc 20040
 gggcctccat cgccttcacc agcatcaacc tctacgccac cttcttcccc atggcgacac 20100
 acaccgcctc cagctcagag gccatgctgc gcaacgacac caacgaccag tccttcaacg 20160
 actacctctc gggggccaac atgctctacc ccatcccggc caacgccacc aacgtgcccc 20220
 tctccatccc ctgcgcgaac tgggcgcgct tccgcggctg gtccttcacg cgcctcaaga 20280
 cccgcgagac gccctcgtc ggctccgggt tcgacctcta cttcgtctac tcggggtcca 20340
 tccctacctc cgacggcacc ttctacctca accacacctt caagaagggt tccatcacct 20400
 tcgactctc cgtcagctgg cccggccaac accgcctct gacgccaac gagtgcgaaa 20460
 tcaagcgac cgtcgacgga gaggggtaca acgtggcccc gtgcaacatg accaaggact 20520
 ggttctggt ccagatgctg gccactaca acatcggtta ccagggttc tacgtgcccg 20580
 agggctacaa ggaccgcatg tactccttct tccgcaactt ccagcccatg agcgcgcagg 20640
 tcgtggacga ggtcaactac aaggactacc aggccgtcac cctggcctac cagcacaaca 20700
 actcgggctt cgtcggctac ctgcgcgcca ccatgcgcca gggccagccc taccgccca 20760
 actacccta cccgtcctc ggcaagagcg ccgtcgccag cgtcaccag aaaaagttcc 20820
 tctgcgaccg ggtcatgtgg cgcctccct tctccagcaa cttcatgtcc atgggcgcgc 20880
 tcaccgacct cggccagaac atgctctacg ccaactccgc ccacgcgcta gacatgaatt 20940
 tcgaagtga ccccatggat gagtccacct ttctctatgt tgtcttcgaa gtcttcgacg 21000
 tcgtccgagt gcaccagccc caccgcggcg tcacgaggc cgtctacctg cgcacgcct 21060
 tctcggccgg caacgccacc acctaaacct cttgcttctt gcaagatgac ggcttcgcgc 21120
 ggctccggcg agcaggagct caggggcctc ctccgcgacc tgggctgcgg gccctgcttc 21180
 ctgggcacct tcgacaagcg cttcccggga ttcattggcc cgcacaagct ggcttcgcgc 21240
 atcgtcaaca cggccggccg cgagaccggg ggcgagcact ggctggcctt cgcctggaac 21300
 ccgcgtccc acacctgcta cctcttcgac cccttcgggt tctcggacga gcgcctcaag 21360
 cagatctacc agttcgagta cgagggcctg ctgcgtcgca gcgcctggc caccgaggac 21420
 cgctgcgtca cctggaaaa gtccaccag accgtgcagg gtccgcgctc ggccgcctgc 21480
 gggctcttct gctgcatgtt cctgcacgcc ttcgtgcact gggccgaccg ccccatggac 21540
 aagaacccca ccatgaactt gctgacgggg gtgccaacg gcatgctcca gtcgccccag 21600

gtggaaccca ccctgcgcgc caaccaggag gcgctctacc gcttctctcaa cgccactcc 21660
 gcctactttc gctcccaccg cgcgcgcacg gagaaggcca ccgccttcga ccgcatgaat 21720
 caagacatgt aatccggtgt gtgtatgtga atgctttatt catcataata aacagcacat 21780
 gtttatgcca ctttctctga ggctctgact ttatttagaa atcgaagggg ttctgcccggc 21840
 tctcggcatg gcccgcgggc agggatacgt tgcggaactg gtacttgggc agccacttga 21900
 actcggggat cagcagcttc ggcacgggga ggtcggggaa cgagtcgctc cacagcttgc 21960
 gcgtgagttg cagggcgccc agcaggtcgg gcgcggagat cttgaaatcg cagttgggac 22020
 ccgcgttctg cgcgcgagag ttacggtaca cggggttgca gcactggaac accatcaggg 22080
 cggggtgctt cagcctcgcc agcaccgtcg cgtcgggtgat gccctccacg tccagatcct 22140
 cggcgttggc catcccgaag ggggtcatct tgcaggtctg ccgccccatg ctgggcacgc 22200
 agccgggctt gtggttgcaa tcgcagtgcg gggggatcag catcatctgg gcctgctcgg 22260
 agctcatgcc cgggtacatg gccttcacga aagcctccag ctggcggaag gcctgctcgg 22320
 ccttgccgcc ctccgtgaag aagaccccgc aggaacttgc agagaactgg ttgggtggcg 22380
 agccagcgtc gtgcacgcag cagcgcgcgt cgttgttggc cagctgcacc acgctgcgcc 22440
 cccagcgggt ctgggtgatc ttggcccggg cggggttctc cttcagcgcg cgctgcccg 22500
 tctcgcctgc cacatccatc tcgatcgtgt gctccttctg gatcatcacg gtcccgtgca 22560
 ggcaccgcag cttgccctcg gcctcgggtg acccgtgcag ccacagcgcg cagccgggtg 22620
 tctcccagtt cttgtgggcg atctgggagt gcgagtgac gaagccctgc aggaagcggc 22680
 ccatcatcgt ggtcagggtc ttgttgctgg tgaaggtcag cggaatgccg cgggtgctcct 22740
 cgttcacata cagggtggcag atacggcggt acacctcgcc ctgctcgggc atcagctgga 22800
 aggcggactt caggctcgtc tccacgcggt accggtccat cagcagcgtc atcacttcca 22860
 tgcccttctc ccaggccgaa acgatcggca ggctcagggg gttcttcacc gttgtcatct 22920
 tagtcgcgcg cgcggaagtc aggggggtcgt tctcgtccag ggtctcaaac actcgttgc 22980
 cgtccttctc ggtgatgcgc acggggggaa agctgaagcc cacggccgcc agctcctcct 23040
 cggcctgctt ttcgtcctcg ctgtcctggc tgatgtcttg caaaggcaca tgcttgggtc 23100
 tgcgggggtt ctttttgggc ggcagaggcg gcggcggaag cgtgctgggc gagcgcgagt 23160
 tctcgtcac cagactatt tcttctcctt ggccgtcgtc cgagaccacg cggcggtagg 23220
 catgcctctt ctggggcaga ggcggaggcg acgggctctc gcggttcggc gggcggctgg 23280
 cagagccctt tccgcgttcg ggggtgcgct cctggcggcg ctgctctgac tgacttcctc 23340

cgcgcccgcc cattgtgttc tccatgggag caagcatgga gactcagcca tcgtcgccaa 23400
 catcgccatc tgcccccgcc gceggcgacg agaaccagca gcagcagaat gaaagcttaa 23460
 ccgccccgcc gccagcccc acctccgacg ccgcagcccc agacatgcaa gagatggagg 23520
 aatccatcga gattgacctg ggctacgtga cggcccgga gcacgaggag gagctggcag 23580
 cgcgcttttc agccccggaa gagaaccacc aagagcagcc agagcaggaa gcagagagcg 23640
 agcagaacca ggctgggctc gagcatggcg actacctgag cggggcagag gacgtgctca 23700
 tcaagcatct ggcccgccaa tgcatcatcg tcaaggacgc gctgctcgac cgcgcccagg 23760
 tgccccctcag cgtggcgagg ctccgcccgc cctacgagcg caacctcttc tcgcccgcg 23820
 tgccccccaa gcgccagccc aacggcacct gcgagccaa cccgcgcctc aacttctacc 23880
 cggctcttcg gggtggcgag gccctggcca cctaccacct ctttttcaag aaccaaagga 23940
 tccccgtctc ctgcccgcgc aaccgcaccc gcgccgacgc cctgctcaac ctgggccccg 24000
 gcgcccgcct acctgatata gcctccttgg aagagggtcc caagatcttc gagggtcttg 24060
 gcagcgacga gactcggggc gcgaacgctc tgcaaggaa cggagaggag catgagcacc 24120
 acagcgccct ggtggagttg gaaggcgaca acgcgcgcct ggcggtcctc aagcgcacgg 24180
 tcgagctgac ccacttcgcc taccggcgcc tcaacctgcc cccaaggte atgagcgccg 24240
 tcatggacca ggtgctcatc aagcggcct cgccccctc ggaggaggag atgcaggacc 24300
 ccgagagctc ggacgagggc aagcccgtgg tcagcgacga gcagctggcg cgtggctgg 24360
 gagcgagtag cccccccag agcctggaag agcggcgcaa gctcatgatg gccgtggctc 24420
 tgggtaccgt ggagctggag tgtctgcgcc gcttcttcgc cgacgcggag accctgcgca 24480
 aggtcgagga gaacctgcac tactcttca gacacgggtt cgtgcgccag gcctgcaaga 24540
 tctccaacgt ggagctgacc aacctggtct cctacatggg catcctgcac gagaaccgcc 24600
 tggggcagaa cgtgctgcac accaccctgc gcggggaggc ccgcccgcac tacatccgcg 24660
 actgctcta cctgtacctc tgceacacct ggacgacggg catgggcgtg tggcagcagt 24720
 gcctggagga gcagaacctg aaagagctct gcaagctcct gcagaagaac ctcaaggccc 24780
 tgtggaccgg gttcgacgag cgcaccaccg ccgcggacct ggccgacctc atcttccccg 24840
 agcgctgcg gctgacgctg cgcaacgggc tgcccgaact tatgagccaa agcatgttgc 24900
 aaaactttcg ctctttcacc ctcgaaagct ccgggatcct gccgctcacc tgcctcgcg 24960
 tgccctcgga ctctgtgcg ctgaccttc gcgagtgcgc cccgcccgtc tggagccact 25020

gctacctgct	gcgccctggcc	aactacctgg	cctaccactc	ggacgtgatc	gaggacgtca	25080
gcggcgaggg	cctgctcgag	tgccactgcc	gctgcaacct	ctgcacgccg	caccgctccc	25140
tggcctgcaa	ccccagctg	ctgagcgaga	cccagatcat	cggcaccttc	gagttgcaag	25200
gccccggcga	gggcaagggg	ggtctgaaac	tcaccccggg	gctgtggacc	tggcctact	25260
tgcgcaagtt	cgtgcccag	gactaccatc	ccttcgagat	caggttctac	gaggaccaat	25320
cccagccgcc	caaggccgag	ctgtcggcct	gcgtcatcac	ccagggggcc	atcctggccc	25380
aattgcaagc	catccagaaa	tcccgcgaag	aatttctgct	gaaaaagggc	caagggtct	25440
acttgacccc	ccagaccgga	gaggagctca	accccagctt	ccccaggat	gccccgagga	25500
agcagcaaga	agctgaaagt	ggagctgccg	ccgccgccgg	aggatttga	ggaagactgg	25560
gagagcagtc	aggcagagga	ggaggagatg	gaagactggg	acagcactca	ggcagaggag	25620
gacagcctgc	aagacagtct	ggaggaggaa	gacgaggtgg	aggaggcaga	ggaagaagca	25680
gccgccgcca	gaccgtcgtc	ctcggcgagg	gaggagaaag	caagcagcac	ggataccatc	25740
tccgctccgg	gtcggggctg	cggcgggcgg	gcccacagta	gatgggacga	gaccggggcg	25800
ttcccgaacc	ccaccacca	gaccggtaag	aaggagcggc	agggatacaa	gtcctggcgg	25860
gggcacaaaa	acgccatcgt	ctcctgcttg	caagcctgcg	ggggcaacat	ctccttcacc	25920
cggcgctacc	tgctcttcca	ccgcgggggtg	aacttcccc	gcaacatctt	gcattactac	25980
cgtcacctcc	acagccccta	ctactgtttc	caagaagagg	cagaaaccca	gcagcagcag	26040
cagcagcaga	aaaccagcgg	cagcagctag	aaaatccaca	gcggcggcag	gtggactgag	26100
gatcgcgcg	aacgagccgg	cgcagaccgg	ggagctgagg	aaccggatct	ttcccaccct	26160
ctatgccatc	ttccagcaga	gtcgggggca	agagcaggaa	ctgaaagtca	agaaccgttc	26220
tctcgctcgt	ctcaccgcga	gttgtctgta	tcacaagagc	gaagaccaac	ttcagcgcac	26280
tctcgaggac	gccgaggctc	tcttcaacaa	gtactgcgcg	ctcactctta	aagagtagcc	26340
cgcgcccgcc	cacacacgga	aaaaggcggg	aattacgtca	ccacctgcgc	ccttcgcccc	26400
accatcatca	tgagcaaaga	gatteccacg	cettacatgt	ggagctacca	gccccagatg	26460
ggcctggccg	ccggcgccgc	ccaggactac	tccacccgca	tgaactggct	cagtgcgggg	26520
cccgcgatga	tctcacgggt	gaatgacatc	cgcgcccacc	gaaaccagat	actcctagaa	26580
cagtcagcga	tcaccgccac	gccccgccat	caccttaatc	cgcgtaattg	gccccgccgc	26640
ctgggtgtacc	aggaaattcc	ccagcccacg	accgtactac	ttccgcgaga	cgcgcaggcc	26700
gaagtccagc	tgactaactc	aggtgtccag	ctggccggcg	gcgcgcacct	gtgtcgtcac	26760

cgccccgctc	aggggtataaa	gcggctggtg	atccgaggca	gaggcacaca	gctcaacgac	26820
gaggtggtga	gctcttcgct	gggtctgcga	cctgacggag	tcttccaact	cgccggatcg	26880
gggagatctt	ccttcacgcc	tcgtcaggcc	gtcctgactt	tggagagtcc	gtcctcgag	26940
ccccgctcgg	gtggcatcgg	cactctccag	ttcgtggagg	agttcactcc	ctcgggtctac	27000
ttcaaccctt	tctccggctc	ccccggccac	taccgggacg	agttcatccc	gaacttcgac	27060
gccatcagcg	agtccgtgga	cggctacgat	tgaatgtccc	atgggtggcgc	ggctgacctt	27120
gctcggcttc	gacacctgga	ccactgccgc	cgcttcgctt	gcttcgctcg	ggatctcgcc	27180
gagtttgctt	actttgagct	gcccggaggag	cacctcagg	gcccggccca	cggagtgcgg	27240
atcgctcgctg	aagggggtct	cgactccac	ctgcttcgga	tcttcagcca	gcgtccgate	27300
ctggccgagc	gcgagcaagg	acagaccctt	ctgaccctgt	actgcatctg	caaccaacccc	27360
ggcctgcatg	aaagtctttg	ttgtctgctg	tgtactgagt	ataataaaag	ctgagatcag	27420
cgactactcc	ggacttcogt	gtgttctctg	tatcaaccag	tccctgttct	tcaccgggaa	27480
cgagaccgag	ctccagctcc	agtgttaagc	ccacaagaag	tacctcacct	ggctgttcca	27540
gggctctccg	atcgccgttg	tcaaccactg	cgacaacgac	ggagtccctg	tgagcggccc	27600
tgccaacctt	actttttcca	cccgcagaag	caagctccag	ctcttccaac	ccttcctccc	27660
cgggacctat	cagtgcgtct	cgggaccttg	ccatcacacc	ttccacctga	tcccgaaatc	27720
cacagcgctg	ctccccgcta	ctaacaacca	aactaccac	caacgccacc	gtcgcgacct	27780
ttcctctggg	tctaatacca	ctaccggagg	tgagctccga	ggtcgaccaa	cctctgggat	27840
ttactacggc	ccctggggagg	tggtaggggt	aatagcgcta	ggcctagttg	cgggtgggct	27900
tttggctctc	tgctacctat	acctcccttg	ctgttcgtac	ttagtgggtg	tgtgttgctg	27960
gtttaagaaa	tggggaagat	caccctagtg	agctcgggtg	tgctgggtgg	gggtgggtgctt	28020
tcgattgtgg	gactggggcg	cgcggctgta	gtgaaggaga	aggccgatcc	ctgcttgcat	28080
ttcaatccc	acaaatgcc	gctgagtttt	cagcccgatg	gcaatcgggtg	cgcgggtgctg	28140
atcaagtgcg	gatgggaatg	cgagaacgtg	agaatcgagt	acaataacaa	gactcggaac	28200
aatactctcg	cgtccgtgtg	gcagcccggg	gaccccgagt	ggtacaccgt	ctctgtcccc	28260
gggtgctgacg	gctccccg	caccgtgaat	aatactttca	tttttgcgca	catgtgagac	28320
acgggtcatgt	ggatgagcaa	gcagtacgat	atgtggcccc	ccacgaagga	gaacatcggtg	28380
gtcttctcca	tcgcttacag	cgtgtgcacg	gcgctaata	ccgctatcgt	gtgcctgagc	28440

attcacatgc tcacgctat tggccccaga aataatgccg aaaaagaaaa acagccataa 28500
 caggtttttt cacacacctt ttccagacca tggcctctgt taaatTTTTTg cttttatttg 28560
 ccagtctcat tgccgtcatt catggaatga gtaatgagaa aattactatt tacactggca 28620
 ctaatcacac attgaaaggc ccagaaaaag ccacagaagt ttcattggtat tgttatttta 28680
 atgaatcaga tgtatctact gaactctgtg gaaacaataa caaaaaaaat gagagcatta 28740
 ctctcatcaa gtttcaatgt ggatctgact taaccctaata taacatcact agagactatg 28800
 taggtatgta ttatggaact acagcaggca ttccggacat ggaattttat caagtttctg 28860
 tgtctgaacc caccacgcct agaatgacca caaccacaaa aactacacct gttaccacta 28920
 tacagctcac taccaatggc ttctttgcca tgcttcaagt ggctgaaaat agcaccagca 28980
 ttcaaccac cccaccagc gaggaattc ccagatccat gattggcatt attggtgctg 29040
 tagtggtgtg catgttgatc atcgcttgt gcatggtgta ctatgccttc tgctacagaa 29100
 agcacagact gaacgacaag ctggaacact tactaagtgt tgaattttta ttttttagaa 29160
 ccatgaagat cctaggcctt ttagtttttt ctatcattac ctctgctcta tgcaattctg 29220
 acaatgagga cgttactgtc gttgtcggat caaattatc actaaaaggc ccagcaaaag 29280
 gtatgcttcc gtggtattgt tgggtcggaa ctgacgagca acagacagaa ctttgcaatg 29340
 ctcaaaaagg caaaacctca aattctaaaa tctctaatta tcaatgcaat ggcactgact 29400
 tagtattgct caatgtcacg aaagcatatg ctggcagtta cacctgccct ggagatgatg 29460
 ccgacaatat gatTTTTTt aaagtggaag tgggtgatcc cactactcca ccgccacca 29520
 ccacaactac tcataccaca cacacagaac aaacaccaga ggcagcagaa gcagagttgg 29580
 ccttcagggt tcacggagat tcccttgctg tcaatacccc tacacccgat cagcgggtgtc 29640
 cggggctgct cgtcagcggc attgtcgggtg tgctttcggg attagcagtc ataactatct 29700
 gcatgttcat ttttgcttgc tgctatagaa ggctttaccg acaaaaatca gaccactgc 29760
 tgaacctcta tgtttaattt ttccagagc catgaaggca gttagcgtc tagtttttg 29820
 ttctttgatt ggcattgttt ttagtgctgg gtttttgaat aatcttacca ttatgaagg 29880
 tgagaatgcc actctagtgg gcatcagtgg tcaaatgtc agctggctaa aataccatct 29940
 agatgggtgg aaagacattt gcgattggaa tgcactgtg tatacatgta atggagttaa 30000
 cctcaccatt actaatgcca cccaagatca gaatggtagg tttaagggcc agagtttcac 30060
 tagaaataat gggatgaat cccaatacat gtttatctat gacgtcactg tcacagaaa 30120
 tgagactgcc accaccacac agatgccac tacacacagt tctaccacta ctacatgca 30180

aaccacacag acaaccacta catcaactca gcatatgacc accactacag cagcaaagcc 30240
aagtagtgca gcgcctcage cccaggcttt ggctttgaaa gctgcacaac ctagtacaac 30300
tactaggacc aatgagcaga ctactgaatt tttgtccact gtcgagagcc acaccacagc 30360
tacctccagt gccttctcta gcacggccaa tctctcctcg ctttcctcta caccatcag 30420
tcccgtact actcccaccc cagctcttct ccccaactcc ctgaagcaaa ctgaggacag 30480
cggcatgcaa tggcagatca ccttgcctcat tgtgatcggg ttggtcatcc tggcgtgtt 30540
gctctactac atcttctgcc gccgcattcc caacgcgcac cgcaaaccgg cctacaagcc 30600
catcgttate gggcagccgg agcgcgttca ggtggaaggg ggtctaagga atcttctctt 30660
ctcttttaca gtatggtgat tgaactatga ttcctagaca attcttgatc actattctta 30720
tctgcctcct ccaagtctgt gccacccctcg ctctgggtggc caacgccagt ccagactgta 30780
ttggggccctt cgcctcctac gtgctctttg ccttcattcac ctgcatctgc tgcgttagca 30840
tagtctgcct gcttatcacc ttcttccagt tcattgactg gatctttgtg cgcctgcct 30900
acctgcgcca ccacccccag taccgcgacc agcagtgggc gcggctgctc aggctcctct 30960
gataagcatg cgggctctgc tacttctcgc gcttctgctg ttagtgetcc cccgccccgt 31020
cgacccccgg tccccactc agtccccga agaggtccgc aaatgcaaat tccaagaacc 31080
ctggaattc ctcaaagtct accgccaaaa atcagacatg cttcccagct ggatcatgat 31140
cattgggata gtgaacattc tggcctgcac cctcatctcc tttgtgattt acccctgctt 31200
tgactttggt tggaaactgc cagaggcgct ctatctccc cctgaacctg acacaccacc 31260
acagcaacct caggcacacg cactaccacc accacagcct aggccacaat acatgcccac 31320
attagactat gaggcgagc cacaggcacc catgctcccc gctattagtt atttcaatct 31380
aaccggcgga gatgactgac ccactggcca acaacaacgt caacgacctt ctctggaca 31440
tggacggccg cgcctcggag cagcgactcg cccaacttcg cattcgccag cagcaggaga 31500
gagccgtcaa ggagctgcag gacggcatag ccatccacca gtgcaagaaa ggcatcttct 31560
gcctggtgaa acaggccaag atctcctacg aggtcacccc gaccgacct cgcctctcct 31620
acgagctcct gcagcagcgc cagaagttca cctgcctggg cggagtcaac cccatcgta 31680
tcaccagca gtcgggcgat accaaggggt gcatccactg ctctgcgac tccccgact 31740
gcgtccacac tctgatcaag accctctgcg gcctccgga cctcctcccc atgaactaat 31800
caccctta tccagtgaat taaatatcat attgatgatg atttaaataa aaaataatca 31860

tttgatttga	aataaagata	caatcatatt	gatgatttga	gttttaaaaa	ataaagaatc	31920
acttacttga	aatctgatac	caggtctctg	tccatgtttt	ctgccaacac	cacctcactc	31980
ccctcttccc	agctctggta	ctgcagaccc	cgcggggctg	caaacttcct	ccacacgctg	32040
aaggggatgt	caaattcctc	ctgtccctca	atcttcattt	tatcttctat	cagatgtcca	32100
aaaagcgct	cgggtggat	gatgacttcg	accccgctc	cccctacgat	gcagacaacg	32160
caccgacgct	gcccttcac	aacccccct	tcgtctcttc	agatggattc	caagagaagc	32220
ccctgggggt	gctgtccctg	cgactggctg	accccgtcac	caccaagaac	ggggaaatca	32280
ccctcaagct	gggagagggg	gtggacctcg	actcctcggg	aaaactcctc	tccaacacgg	32340
ccaccaaggc	cgccgccct	ctcagttttt	ccaacaacac	catttcctct	aacatggata	32400
ccctctttta	taccaaagat	ggaaaattat	ccttacaagt	ttctccaccg	ttaaaccatat	32460
taaaatcaac	cattctgaac	acattagctg	tagcttatgg	atcaggttta	ggactgagtg	32520
gtggcactgc	tcttgcagta	cagttggcct	ctccactcac	ttttgatgaa	aaaggaaata	32580
ttaaaattaa	cctagccagt	ggtccattaa	cagttgatgc	aagtcgactt	agtatcaact	32640
gcaaaagagg	ggtcactgtc	actacctcag	gagatgcaat	tgaaagcaac	ataagctggc	32700
ctaaagggtat	aagatttgaa	ggtaatggca	tagctgcaaa	cattggcaga	ggattggaat	32760
ttggaaccac	tagtacagag	actgatgtca	cagatgcata	cccaattcaa	gttaaattgg	32820
gtactggcct	tacctttgac	agtacaggcg	ccattgttgc	ttggaacaaa	gaggatgata	32880
aacttacatt	atggaccaca	gccgaccct	cgccaaattg	caaaatatac	tctgaaaaag	32940
atgccaaact	cacactttgc	ttgacaaagt	gtggaagtca	aattctgggt	actgtgactg	33000
tattggcagt	gaataatgga	agtctcaacc	caatcacaaa	cacagtaagc	actgcactcg	33060
tctccctcaa	gtttgatgca	agtggagttt	tgctaagcag	ctccacatta	gacaaagaat	33120
attggaactt	cagaaaggga	gatgttacac	ctgctgagcc	ctataactaat	gctatagggt	33180
ttatgcctaa	cataaaggcc	tatcctaaaa	acacatctgc	agcttcaaaa	agccatattg	33240
tcagtcaagt	ttatctcaat	ggggatgagg	ccaaaccact	gatgctgatt	attactttta	33300
atgaaactga	ggatgcaact	tgcacctaca	gtatcacttt	tcaatggaaa	tgggatagta	33360
ctaagtacac	aggtgaaaca	cttgctacca	gctccttcac	cttctcctac	atcgcccaag	33420
aatgaacact	gtatcccacc	ctgcatgcca	acccttccca	ccccactctg	tctatggaaa	33480
aaactctgaa	gcacaaaata	aaataaagtt	caagtgtttt	attgattcaa	cagttttaca	33540
ggattcgagc	agttattttt	cctccaccct	cccaggacat	ggaatacacc	accctctccc	33600

cccgcacagc	cttgaacatc	tgaatgccat	tggtgatgga	catgcttttg	gtctccacgt	33660
tccacacagt	ttcagagcga	gccagtctcg	ggtcgggtcag	ggagatgaaa	ccctccgggc	33720
actcccgc	ctgcacctca	cagctcaaca	gctgaggatt	gtcctcgggtg	gtcgggatca	33780
cggttatctg	gaagaagcag	aagagcggcg	gtgggaatca	tagtccgcga	acgggatcgg	33840
cgggtggtgt	cgcatacaggc	cccgacagcag	tcgctgccgc	cgccgctccg	tcaagctgct	33900
gctcaggggg	tccgggtcca	gggactccct	cagcatgatg	cccacggccc	tcagcatcag	33960
tcgtctggtg	cggcggggcg	agcagcgc	gctggtctcg	ctcaggctcg	tgcagtacgt	34020
gcaacacagg	accaccagg	tgttcaacag	tccatagtcc	aacacgctcc	agccgaaact	34080
catcgcgga	aggatgctac	ccacgtggcc	gtcgtaccag	atcctcagg	aatcaagt	34140
gcgctccctc	cagaacacgc	tgcccacgta	catgatctcc	ttgggcatgt	ggcgggtcac	34200
cacctcccgg	taccacatca	ccctctgggt	gaacatgcag	ccccggatga	tcctgcggaa	34260
ccacagggcc	agcaccgccc	cggccgcat	gcagcgaaga	gaccccggt	cccggcaatg	34320
gcaatggagg	acccaccgct	cgtaccgtg	gatcatctgg	gagctgaaca	agtctatgtt	34380
ggcacagcac	aggcatatgc	tcatgcatct	cttcagcact	ctcagctcct	cgggggtcaa	34440
aaccatatcc	cagggcacgg	ggaactcttg	caggacagcg	aaccccgag	aacagggcaa	34500
tcctcgaca	taacttacat	tgtgcatgga	cagggtatcg	caatcaggca	gcaccgggtg	34560
atcctccacc	agagaagcgc	gggtctcgg	ctcctcacag	cgtggtaagg	gggcccggccg	34620
ataggggtga	tggcgggacg	cggctgatcg	tgttcgcgac	cgtgtcatga	tgcagttgct	34680
ttcggacatt	ttcgtacttg	ctgtagcaga	acctggctcg	ggcgtgcac	accgatcgcc	34740
ggcggcggtc	ccggcgcttg	gaacgctcgg	tgttgaaatt	gtaaaacagc	cactctctca	34800
gaccgtgcag	cagatctagg	gcctcaggag	tgatgaagat	cccatcatgc	ctgatagctc	34860
tgatcacatc	gaccaccgtg	gaatgggcca	gaccagcca	gatgatgcaa	ttttgttggg	34920
tttcggtgac	ggcgggggag	ggaagaacag	gaagaacat	gattaacttt	taatccaaac	34980
ggtctcggag	cacttcaaaa	tgaaggctcg	ggagatggca	cctctcgccc	ccgctgtgtt	35040
ggtggaaaat	aacagccagg	tcaaagggtga	tacggttctc	gagatgttcc	acgggtggctt	35100
ccagcaaagc	ctccacgcgc	acatccagaa	acaagacaat	agcgaaagcg	ggagggttct	35160
ctaattcctc	aatcatcatg	ttacactcct	gcaccatccc	cagataattt	tcatttttcc	35220
agccttgaat	gattcgaact	agttcctgag	gtaaatccaa	gccagccatg	ataaagagct	35280

cgcgagagc gccctccacc ggcattctta agcacaccct cataattcca agatattctg 35340
 ctctgtgttc acctgcagca gattgacaag cggaatatca aaatctctgc cgcgatccct 35400
 aagctcctcc ctcagcaata actgtaagta ctctttcata tcctctccga aatttttagc 35460
 cataggacca ccaggaataa gattagggca agccacagta cagataaacc gaagtcctcc 35520
 ccagtgcgca ttgccaaatg caagactgct ataagcatgc tggctagacc cgggtgatatc 35580
 ttccagataa ctggacagaa aatcaccag gcaattttta agaaaatcaa caaaagaaaa 35640
 atcctccagg tgcacgttta gagcctcggg aacaacgatg aagtaaagtc aagcgggtgcg 35700
 ttccagcatg gttagtttagc tgatctgtaa aaaacaaaaa ataaaacatt aaaccatgct 35760
 agcctggcga acaggtgggt aaatcgttct ctccagcacc aggcaggcca cgggggtctcc 35820
 ggcgcgaccc tcgtaaaaat tgcgctatg attgaaaacc atcacagaga gacgttccccg 35880
 gtggccggcg tgaatgattc gacaagatga atacaccccc ggaacattgg cgteccgcgag 35940
 tgaaaaaaag cgcccagga agcaataagg cactacaatg ctcagtctca agtcacgaa 36000
 agcgatgcc a tgcggatgaa gcacaaaatc ctcagggtgcg tacaaaatgt aattactccc 36060
 ctctgcaca ggcagcgaag ccccgatcc ctccagatac acatacaaag cctcagcgtc 36120
 catagcttac cgagcagcag cacacaacag gcgcaagagt cagagaaagg ctgagctcta 36180
 acctgtccac ccgctctctg ctcaatatat agcccagatc tacactgacg taaaggccaa 36240
 agtctaaaaa taccgcgcaa ataatcacac acgcccagca cagcccaga aaccgggtgac 36300
 aactcaaaa aaatacgcg acttctcaa acgcccacac tgcgctcatt tccgggttcc 36360
 cagctacgt catcggaatt cgactttcaa attcgcgca cggtaaaaaa cgtcacccgc 36420
 cccgcccta acggtcgccc gtctctcggc caatcacctt cctccctccc caaattcaaa 36480
 cagctcattt gcatattaac gcgcacaaa agtttgaggt atattattga tgatg 36535

<210> 4
 <211> 34264
 <212> DNA
 <213> simian adenovirus SV-1

<400> 4
 tccttattct ggaaacgtgc caatatgata atgagcgggg aggagcgagg cggggccggg 60
 gtgacgtgcg gtgacgtggg gtgacgcggg gtggcgcgag ggcggggcgg gagtggggag 120
 gcgcttagtt ttacgtatg cggaaggagg ttttataacc gaagttgggt aatttgggag 180
 tataacttgta agttttgtgt aatttggcgc gaaaaccggg taatgaggaa gttgagggtta 240

atatgtactt tttatgactg ggcggaattt ctgctgatca gcagtgaact ttgggcgctg	300
acggggaggt ttcgctacgt ggcagtagca cgagaaggct caaagggtccc atttattgta	360
ctcctcagcg ttttcgctgg gtatttaaac gctgtcagat catcaagagg ccactcttga	420
gtgccggcga gtagagtttt ctccctcgcg ctgccgcgat gaggctgggt cccgagatgt	480
acgggtgttt ctgcagcgag acggccccga actcagatga gctgcttaac acagatctgc	540
tggatgttcc caactcgct gtggcttcgc ctccgtcgct tcatgatctt ttcgatgtgg	600
aagtggatcc accgcaagat cccaacgagg acgcggtaaa cagtatgttc cctgaatgtc	660
tgtttgaggc ggctgaggag ggttctcaca gcagtgaaga gagcagacgg ggagaggaac	720
tggacttgaa atgctacgag gaatgtctgc cttctagcga ttctgaaacg gaacagacag	780
ggggagacgg ctgtgagtcg gcaatgaaaa atgaacttgt attagactgt ccagaacatc	840
ctggctcatgg ctgccgtgcc tgtgcttttc atagaaatgc cagcggaaat cctgagactc	900
tatgtgctct gtgttatctg cgccttacca gcgattttgt atacagtaag taaagtgttt	960
tcattggcgt acggtagggg attcgttgaa gtgctttgtg acttattatg tgtcattatt	1020
tctaggtgac gtgtccgacg tggaaaggga aggagataga tcaggggctg ctaattctcc	1080
ttgcactttg ggggctgtgg ttccagttgg cttttttaa cagagtgggt gaggagaacg	1140
agccggagga gaccgagaat ctgagagccg gcctggaccc tccagtggaa gactaggtgc	1200
tgaggatgat cctgaagagg ggactagtgg ggggtgctagg aaaaagcaaa aaactgagcc	1260
tgaacctaga aactttttga atgagttgac tgtaagccta atgaatcggc agcgtcctga	1320
gacggtgttt tggactgagt tggaggatga gttcaagaag ggggaattaa acctcttgta	1380
caagtatggg tttgagcagt tgaaaactca ctggttggag ccgtgggagg atatggaaat	1440
ggctctagac acctttgcta aagtggctct gcggccggat aaagtgtaca ctattcgccg	1500
cactgttaat ataaaaaaga gtgtttatgt tatcgcccat ggagctctgg tgcaggtgca	1560
gacccagac cgggtggctt tcaattgcgg catgcagagt ttgggccccg gggatagagg	1620
tttgaatgga gttacatttc aaaatgtcag gtttactggg gatgatttta atggctctgt	1680
gtttgtgact agcaccagc taacctcca cgggtgttac ttttttaact ttaacaatac	1740
atgtgtggag tcatggggtg ggggtgtctc gaggggctgc agttttcatg gttgctggaa	1800
ggcgggtgtg ggaagaatta aaagtgtcat gtctgtgaag aaatgcatat ttgaacgctg	1860
tgtgatagct ctagcagtag aggggtacgg acggatcagg aataacgccg catctgagaa	1920
tggatgtttt cttttgctga aaggtaacgc cagcgttaag cataatatga tttgcggcag	1980

cggcctgtgc	ccctcgcagc	tcttaacttg	cgagatgga	aactgtcaca	ccttgcgcac	2040
cgtgcacata	gtgtcccact	cgcgcgcac	ctggccaaca	tttgagcaca	atatgtcat	2100
gcgttgcgcc	gttcacctag	gtgctagacg	cggcgtgttt	atgccttacc	aatgtaactt	2160
tagtcatact	aagattttgc	tggaaactga	tcccttccct	cgagtatgtt	tcaatggggt	2220
gtttgacatg	tcaatggaac	tttttaaagt	gataagatat	gatgaaacca	agtctcgttg	2280
tcgctcatgt	gaatgcggag	ctaatacttt	gaggttgat	cctgtaaccc	tgaacgttac	2340
cgaggagctg	aggacggacc	accacatgct	gtcttgccctg	cgtaccgact	atgaatccag	2400
cgatgaggag	tgaggtgagg	ggcggagcca	caaagggat	aaaggggcat	gaggggtggg	2460
cgcgggtgtt	caaaatgagc	gggacgacgg	acggcaatgc	gtttgagggg	ggagtgttca	2520
gcccatactt	gacatctcgt	cttccttccct	gggcaggagt	tcgtcagaat	gtagtgggct	2580
ccaccgtgga	cggacggccg	gtcgccctg	caaattccgc	caccctcacc	tatgccaccg	2640
tgggatcatc	gttggacact	gccgcggcag	ctgccgcttc	tgctgccgct	tctactgctc	2700
gcggcatggc	ggctgatttt	ggactatata	accaactggc	cactgcagct	gtggcgcttc	2760
ggctctcgtt	tcaagaagat	gccctgaatg	tgatcttgac	tcgcctggag	atcatgtcac	2820
gtcgctgga	cgaaactggc	gcgcagatat	cccaagctaa	ccccgatacc	gcttcagaat	2880
cttaaaataa	agacaaacaa	atttggtgaa	aagtaaaatg	gctttatttg	ttttttttgg	2940
ctcggtaggc	tcgggtccac	ctgtctcgtt	cgtaaggac	tttggtgatg	ttttccaaaa	3000
cacggtacag	atgggcttgg	atgttcaagt	acatgggcat	gaggccatct	ttggggtgga	3060
gataggacca	ctgaagagcg	tcattgttccg	gggtggtatt	gtaaatcacc	cagtcgtagc	3120
agggtttttg	agcgtggaac	tggaaatagt	ccttcaggag	caggctaattg	gccaaagggt	3180
gacccttagt	gtaggtgttt	acaaagcggg	tgagctggga	gggatgcatg	cggggggaga	3240
tgatatgcat	cttggcttgg	attttgaggt	tagctatgtt	accacccagg	tctctgcggg	3300
ggttcatgtt	atgaaggacc	accagcacgg	tatagccagt	gcatttgggg	aacttgtcat	3360
gcagtttgga	ggggaaggcg	tggaaagaatt	tagatacccc	cttgtgcccc	cctaggtttt	3420
ccatgcactc	atccataata	atggcaatgg	gacccctggc	ggccgcttta	gcaaacacgt	3480
tttgggggtt	ggaaacatca	tagttttgct	ctagagttag	ctcatcatag	gccatcttta	3540
caaagcgggg	taggaggggtg	cccgaactggg	ggatgatagt	tccatctggg	cctggagcgt	3600
agttgccctc	acagatctgc	atctcccagg	ccttaatttc	cgaggggggg	atcatgtcca	3660

cctggggggc	gataaaaaac	acggtttctg	gcggggggtt	aatgagctgg	gtggaaagca	3720
agttacgcaa	cagctgggat	ttgccgcaac	cgggtgggacc	gtagatgacc	ccgatgacgg	3780
gttgacgctg	gtagttcaga	gaggaacagc	tgccgtcggg	gcgcaggagg	ggagctacct	3840
cattcatcat	gcttctgaca	tgtttatttt	cactcactaa	gttttgcaag	agcctctccc	3900
caccagggga	taagagttct	tcagggctgt	tgaagtgttt	cagcgggttc	aggcgtcgg	3960
ccatgggcat	cttttcaagc	gactgacgaa	gcaagtacag	tcgggtcccag	agctcgggtga	4020
cgtgctctat	ggaatctcga	tcacgcagac	ttcttggttt	cgggggttgg	gccgactttc	4080
gctgtagggc	accagccggg	gggcgtccag	ggccgcgagg	gttctgtcct	tcagggtct	4140
cagcgttcgg	gtgaggggtg	tctcgggtgac	ggtgaaggga	tgagccccgg	gctgggcgt	4200
tgcgagggtg	cgcttcaggc	tcattcctgt	ggtgctgaag	cgggcgtcgt	ctccctgtga	4260
gtcgcccaga	tagcaacgaa	gcatgaggtc	gtagctgagg	gactcggccg	cgtgtccctt	4320
ggcgccgacg	tttcccttgg	aaacgtgctg	acatttgggtg	cagtgcagac	acttgagggg	4380
gtagagtttt	ggggccagga	agaccgactc	gggcgagtag	gcgtcggctc	cgcactgagc	4440
gcagacggtc	tcgcactcca	ccagccacgt	gagctcgggt	ttagcgggat	caaaaaccaa	4500
gttgccctcca	ttttttttga	tgcgtttctt	accttgcgctc	tccatgagtc	tgtgtcccg	4560
ttccgtgaca	aaaaggctgt	cgggtatccc	gtagaccgac	ttgagggggc	gatcttccaa	4620
aggtgttccg	aggtcttccg	cgtacaggaa	ctgggaccac	tccgagacaa	aggctcgggt	4680
ccaggctaac	acgaaggagg	cgatctgcga	ggggtatctg	tcgttttcaa	tgagggggtc	4740
caccttttcc	agggtgtgca	gacacaggtc	gtcctcctcc	gcgtccacga	agggtattgg	4800
cttgtaagtg	taggtcacgt	gaccgcaccc	cccccaaggg	gtataaaagg	gggcgtgccc	4860
actctccccg	tcactttctt	ccgcctcgct	gtggaccaga	gccagctgtt	cgggtgagta	4920
ggccctctca	aaagccggca	tgatttcggc	gctcaagttg	tcagtttcta	caaacgagggt	4980
ggatttgata	ttcacgtgcc	ccgcggcgat	gcttttgatg	gtggaggggt	ccatctgac	5040
agaaaacacg	atctttttat	tgtcaagtgt	ggtggcgaaa	gaccgtaga	gggcgttgg	5100
aagcaacttg	gcgatggagc	gcagggtctg	atttttctcc	cgatcgcccc	tctccttggc	5160
ggcgatgttg	agttgcacgt	actcgcgggc	cacgcaccgc	cactcgggga	acacggcgggt	5220
gcgctcgtcg	ggcaggatgc	gcacgcgcc	gccgcgggtg	tgaggggtga	tgaggtccac	5280
gctgggtggc	acctccccgc	ggaggggctc	gttgggtccaa	cacaatcgcc	ccccttttct	5340
ggagcagaac	ggaggcaggg	gatctagcaa	gttggcgggc	ggggggctcg	cgtcgatgg	5400

aaatatgccg ggtagcagaa ttttattaaa ataatcgatt tcggtgtccg tgtcttgcaa	5460
cgcgtcttcc cacttcttca ccgccagggc cctttcgtag ggattcaggg gcggtcccca	5520
gggcatgggg tgggtcaggg ccgaggcgta catgccgcag atgtcgtaca cgtacagggg	5580
ctccctcaac accccgatgt aagtggggta acagcgcccc ccgcggatgc tggctcgac	5640
gtagtcgtac atctcgtgag agggagccat gagcccgtct cccaagtggg tcttgtgggg	5700
tttttcggcc cggtagagga tctgectgaa gatggcgtag gagttggaag agatagtggg	5760
gcgttggaag acgttaaagt tggctccggg cagtcaccag gagtcttgga tgaactgggc	5820
gtaggattcc cggagcttgt ccaccagggc tgcggttacc agcacgtcga gagcgcagta	5880
gtccaacgtc tcgcggacca ggttgtaggc cgtctcttgt ttttctctcc acagttcgcg	5940
attgaggagg tattcctcgc ggtctttcca gtactcttcg gcgggaaatc ctttttcgtc	6000
cgtcggtaa gaacctaaac tgtaaaattc gttcacggct ttgtatggac aacagccttt	6060
ttctaccggc agggcgtaac cttgagcggc ctttctgaga gaggtgtggg tgagggcgaa	6120
ggtgtccgc accatcactt tcaggactcg atgtttgaag tccgtgtcgt cgcaggcgcc	6180
ctgttccac agcgtgaagt cggtcgctt tttctgctg ggattgggga gggcgaatgt	6240
gacgtcgta aagaggatth tcccggcgcg gggcatgaag ttgcgagaga tcctgaaggg	6300
tccgggcacg tccgagcggc tgttgatgac ttgcgcgcc aggacgatct cgtcgaagcc	6360
gttgatgttg tggccacga tgtaaaagtc gataaagcgc ggctgtccct tgagggccgg	6420
cgttttttc aactcctcgt aggtgagaca gtccggcgag gagagacca gctccgccc	6480
ggcccagtcg gagagctgag ggttagccgc gagaaagag ctccacaggt caagggctag	6540
cagagtttgc aagcggtcgc ggaactcgcg aaacttttc cccacggcca tttctcgg	6600
cgtcaccacg tagaaagtgc agggcggtc gttccagacg tcccatcgga gctctagggc	6660
cagctcgag gcttgacgaa cgagggtctc ctgcccag agtgcatga ccagcatgaa	6720
gggtaccaac tgtttccga acgagcccat ccatgtgtag gtttctacgt cgtaggtgac	6780
aaagagccgc tgggtgcgc cgtgggagcc gatcgggaag aagctgatct cctgccacca	6840
gttgaggaa tgggtgttga tgtggtgaaa gtagaagtc ccgcggcgca cagagcattc	6900
gtgctgatgt ttgtaaaagc gaccgcagta gtcgcagcgc tgcacgtct gtatctcctg	6960
aatgagatgc gcttttcgcc cgcgcaccag aaaccggagg gggagttga gacgggggct	7020
tgggtggggc gcatccctt cgccttggcg gtgggagtc gcgtctgcgc cctccttctc	7080

tgggtggacg acggtgggga cgacgacgcc ccgggtgccg caagtccaga tctcdgccac	7140
ggagggggcgc aggcgttgca ggagggggacg cagctgcccc ctgtccaggg agtcgagggc	7200
ggccgcgctg aggtcggcgg gaagcgtttg caagttcact ttcagaagac cggtaagagc	7260
gtgagccagg tgcacatggt acttgatttc caggggggtg ttggaagagg cgtccacggc	7320
gtagaggagg ccgtgtccgc gcggggccac caccgtgccc cgaggagggt ttatctcact	7380
cgtcgagggc gagcgccggg gggtagaggc ggctctgcgc cggggggcag cggaggcagt	7440
ggcacgtttt cgtgaggatt cggcagcggg tgatgacgag cccggagact gctggcgtgg	7500
gcgacgacgc ggcggttgag gtcctggatg tgccgtctct gcgtgaagac caccggcccc	7560
cgggtcctga acctgaaaga gagttccaca gaatcaatgt ctgcatcggt aacggcgggc	7620
tgcttgagga tctcctgtac gtcgcccag ttgtcttgat aggcgatctc ggccatgaac	7680
tgctccactt ctctctcgcg gaggtcgccg tggcccgtc gctccacggg ggcgccagg	7740
tcgttgagga tgcgacgcat gagttgagag aaggcggtga ggccgttctc gttccacacg	7800
cggctgtaca ccacgtttcc gaaggagtcg cgcgctcgca tgaccacctg ggccacgttg	7860
agttccacgt ggcgggcgaa gacggcgtag tttctgaggc gctggaagag gtagttgagc	7920
gtggtggcga tgtgtcgca gacgaagaag tacatgateg agcgccgcag ggtcatctcg	7980
ttgatgtctc cgatggcttc gagacgctcc atggcctcgt agaagtcgac ggccaagtgt	8040
aaaaattggg agttgcgggc ggccaccgtg agttcttctt gcaggaggcg gatgagatcg	8100
gcgaccgtgt cgcgcacctc ctgctcgaaa gcgccccgag gcgcctctgc ttcttctctc	8160
ggctctctct ctctcagggg caccgggttc tccggcagct ctgcgacggg gacggggcgg	8220
cgacgtcgtc gtctgaccgg caggcggtcc acgaagcgt cgatcatttc gccgcgccgg	8280
cgacgcatgg tctcggtgac ggcggtccg ttttcgcgag gtcgcagttc gaagacgccg	8340
ccgcgcagag cgcgccgtg caggggagggt aagtggttag ggccgtcggg cagggaacacg	8400
gcgctgacga tgcattttat caattgctgc gtaggcactc cgtgcaggga tctgagaacg	8460
tcgaggtcga cgggatccga gaacttctct aggaaagcgt ctatccaatc gcagtcgcaa	8520
ggtaagctga ggacggtggg ccgctggggg gcgtccgcgg gcagttggga ggtgatgctg	8580
ctgatgatgt aattaaagta ggcggtcttc aggcggcgga tggtagcgag gaggaccacg	8640
tctttgggccc cggcctgttg aatgcgcagg cgctcgcca tgccccaggc ctcgctctga	8700
cagcgacgca ggtctttgta gtagtcttgc atcagttctt ccaccggaac ctctgcttct	8760
cccctgtctg ccatgcgagt cgagccgaac cccgcagggg gctgcagcaa cgctaggctg	8820

gccacgaccc	tctcggccag	cacggcctgt	tggatctgcg	tgaggggtgg	ctggaagtcg	8880
tccaggtcca	cgaagcgggt	ataggccccc	gtgttgatgg	tgtaggtgca	gttggccatg	8940
acggaccagt	tgacgacttg	catgccgggt	tgggtgatct	ccgtgtactt	gaggcgcgag	9000
tagggcgggg	actcgaacac	gtagtctgtg	catgtgctga	ccagatactg	gtagccaacc	9060
aggaagtggg	gaggcgggtc	tcggtacagg	ggccagccga	ctgtggcggg	ggcgccgggg	9120
gacaggtcgt	ccagcatgag	gcgatggtag	tggtagatgt	agcgggagag	ccaggtgatg	9180
ccggccgagg	tggtcgcggc	cctggtgaat	tcgcggacgc	ggttccagat	gttgccgagg	9240
gggcgaaagc	gctccatggt	gggcacgctc	tgccccgtga	ggcgggcgca	atcttgtagc	9300
ctctagatgg	aaaaaagaca	gggcggtcac	cgactccctt	ccgtagctcg	gggggttaaag	9360
tcgcaagggg	gcggcgggcg	ggaaccccg	ttcgagaccg	gccggatccg	ccgctcccga	9420
tgccctggc	cccgcaccca	cgacgtccgc	gtcgagaccc	agccgcgacg	ctccgccccca	9480
atacggaggg	gagtcttttg	gtgttttttc	gtagatgcat	ccggtgctgc	ggcagatgcg	9540
acctcagacg	cccaccacca	ccgccggggc	ggcagtaaag	ctgagcggag	gcgggtgacag	9600
ggaggaggag	gagctggctt	tagacctgga	agagggagag	gggctggccc	ggctgggagc	9660
gccgtcccca	gagagacacc	ctagggttca	gctcgtgagg	gacgccaggc	aggcttttgt	9720
gccgaagcag	aacctgttta	gggaccgcag	cggtcaggag	gcggaggaga	tgccgcgattg	9780
caggtttcgg	gcgggtagag	agctgagggc	gggcttcgat	cgggagcggc	tcctgagggc	9840
ggaggatttc	gagcccgacg	agcgttctgg	ggtgagcccg	gcccgcgctc	acgtctcggc	9900
ggccaacctg	gtgagcgcgt	acgagcagac	ggtgaacgag	gagcgcaact	tccaaaagag	9960
ctttaacaat	cacgtgagga	ccctgatcgc	gagggaggag	gtgaccatcg	ggctgatgca	10020
tctgtgggac	ttcgtggagg	cctacgtgca	gaaccggggc	agcaaaccctc	tgacggccca	10080
gctgttcctg	atcgtgcagc	acagccgcga	caacgagacg	ttccgcgacg	ccatgttgaa	10140
catcgcggag	cccgaggggc	gctggctctt	ggatctgatt	aacatcctgc	agagcatcgt	10200
ggtgcaggag	aggggcctca	gcttagcggg	caagggtggc	gccattaaact	attcgatgca	10260
gagcctgggg	aagttctacg	ctcgcaagat	ctacaagagc	ccttacgtgc	ccatagacaa	10320
ggagggtgaag	atagacagct	tttacatgcg	catggcgctg	aagggtgctga	cgctgagcga	10380
cgacctcggc	gtgtaccgta	acgacaagat	ccacaaggcg	gtgagcgcca	gccgccggcg	10440
ggagctgagc	gacagggagc	tgatgcacag	cctgcagagg	gcgctggcgg	gcgccgggga	10500

cgaggagcgc gaggettact tcgacatggg agccgatctg cagtggcgtc ccagcgcgcg 10560
 cgccttggag gcggcgggct accccgacga ggaggatcgg gacgatttgg aggaggcagg 10620
 cgagtacgag gacgaagcct gaccgggcag gtgttgtttt agatgcagcg gccggcggac 10680
 gggggccaccg cggatcccg acttttggca tccatgcaga gtcaaccttc gggcgtgacc 10740
 gcctccgatg actgggcggc ggccatggac cgcattatgg cgctgactac ccgcaacccc 10800
 gaggctttta gacagcaacc ccaggccaac cgtttttcgg ccatcttga agcgggtggtg 10860
 ccctcccga ccaaccccac acacgagaaa gtcctgacta tcgtgaacgc cctggtagac 10920
 agcaaggcca tccgccgga cgaggcgggc ttgatttaca acgctctgct ggaacgggtg 10980
 gcgcgctaca acagcactaa cgttcagacc aatctggatc gcctcaccac cgacgtgaag 11040
 gaggcgctgg ctcaagaagg gcggtttctg agggacagca atctgggctc tctggtagca 11100
 ctcaacgcct tcctgagcac gcagcgggc aacgtgcccc gcgggcagga ggactacgtg 11160
 agcttcatca gcgctctgag gctgctggtg tccgaggtgc ccagagcga ggtgtatcag 11220
 tctgggccgg attacttctt ccagacgtcc cgacagggct tgcaaacggt gaacctgact 11280
 caggccttta aaaacttgca aggcattgtg ggcgttaagg ccccggtggg cgatcgagcc 11340
 accatctcca gtctgtgac ccccaacact cgcctgctgc tgctcttgat cgcgccgttc 11400
 accaacagta gcactatcag ccgtgactcg tacctgggtc atctcatcac tttgtaccgc 11460
 gagggcatcg gtcaggctca gatcgacgag cacacatata aggagatcac taacgtgagc 11520
 cgggcccctgg gtcaggaaga taccggcagc ctggaagcca cgttgaactt tttgctaacc 11580
 aaccggaggc aaaaaatacc ctcccagttt acgttaagcg ccgaggagga gaggattctg 11640
 cgatacgtgc agcagtcctg gagtctgtac ttgatgcggg agggcgccac cgcttccacg 11700
 gcttttagaca tgacggctcg gaacatggaa ccgtcctttt actccgccc cgggccgttc 11760
 attaacctgc tgatggacta cttccatcgc gcggccgcca tgaacgggga gtacttcacc 11820
 aatgccatcc tgaatccga ttggatgccc ccgtccggct tctacaccgg cgagtgtgac 11880
 ctgcccgaag ccgacgacgg ctttcttttg gacgacgtgt ccgacagcat tttcacgccg 11940
 ggcaatcgcc gattccagaa gaaggagggc ggagacgagc tccccctctc cagcgtggag 12000
 gcggcctcta ggggagagag tccctttccc agtctgtctt ccgccagcag tggtcgggta 12060
 acgcgcccgc ggttgccggg ggagagcgac tacctgaacg accccttgct gcggccggct 12120
 aggaagaaaa atttcccaa caacgggggtg gaaagcttgg tggataaaat gaatcgttgg 12180
 aagacctacg cccaggagca gcgggagtgg gagacagtc agccgcgacc gctggttccg 12240

ccgcactggc gtcgtcagag agaagacccg gacgactccg cagacgatag tagcgtgttg 12300
 gacctgggag ggagcggagc caaccccttt gtcacttgc aaccaaggg gcgttccagt 12360
 cgctctact aataaaaaag acgcggaaac ttaccagagc catggccaca gcgtgtgtcc 12420
 tttcttctc tctttcttcc tcggcgcgcc agaatgagaa gagcggtag agtcacgccg 12480
 gggcggtatg aggggtccgc cccttcttac gaaagcgtga tgggatcagc gaacgtgccg 12540
 gccacgctgg aggcgcctta cgttctctcc agatacctgg gacctacgga gggcagaaac 12600
 agcatccgtt actccgagct ggcacccctg tacgatacca ccaaggtgta cctgggtggac 12660
 aacaagtcgg cggacatcgc ctccctgaat tatcaaaacg atcacagcaa ttttctgact 12720
 accgtgggtgc agaacaatga cttaccccg acggaggcgg gcacgcagac cattaacttt 12780
 gagcagcgtt cccgctgggg cggtcagctg aaaaccatcc tgcacaccaa catgcccaac 12840
 atcaacgagt tcatgtccac caacaagttc agggccaggc tgatgggttaa aaaggctgaa 12900
 aaccagctc ccgagtacga atggtttgag ttcaccattc ccgagggcaa ctattccgag 12960
 accatgacta tcgatctgat gaacaatgcg atcgtggaca attacctgca agtggggagg 13020
 cagaacgggg tattggaaag cgatatcggc gtaaaatttg ataccagaaa cttccgactg 13080
 ggggtgggatc ccgtgaccaa gctggtgatg ccaggcgtgt acaccaacga ggcttttcac 13140
 ccgacatcg tgctgctgcc ggggtgcggg gtggacttca ctacagagccg tttgagtaac 13200
 ctggttaggga tcagaaagcg ccgccccttc caagagggct ttcagatcat gtatgaggac 13260
 ctggaaggag gtaacattcc aggtttgcta gacgtgccgg cgtatgaaga gagtgttaaa 13320
 caggcggagg cgcagggacg agagattcga ggcgacacct ttgccacgga acctcacgaa 13380
 ctggttaataa aacctctgga acaagacagt aaaaaacgga gttacaacat tatatccggc 13440
 actatgaata ccttgtaccg gagctgggtt ctggcttaca actacgggga tcccgaaaag 13500
 ggagtgagat catggaccat actcaccacc acggacgtga cctgcggctc gcagcaagtg 13560
 tactggctcc tgccggatat gatgcaagac cgggtcacct tccgcccctc caccgaagt 13620
 agcaacttcc cgggtggggg caccgagctg ctgcccgctc atgccaagag cttctacaac 13680
 gaacaggccg tctactcgca actcatcgc cagtccaccg cgcttaccga cgtgttcaat 13740
 cgctttcccg agaaccagat tctgggtgcg cctcccgcct ctaccattac caccgtcagt 13800
 gaaaacgttc ccgccctcac agatcacgga accctgccgc tgcgcagcag tatcagtggg 13860
 gttcagcgcg tgaccatcac cgaöccaga cgtcgaacct gtccctacgt ttacaaagct 13920

cttggcgtag	tggtctcctaa	agtgtctctct	agtcgcacct	tctaaacatg	tccatcctca	13980
tctctcccga	taacaacacc	ggctggggac	tgggctccgg	caagatgtac	ggcggagcca	14040
aaaggcgctc	cagtcagcac	ccagttcgag	ttcggggcca	cttcctgtgct	ccctggggag	14100
cttacaagcg	aggactctcg	ggcgcgaacgg	cggtagacga	taccatagat	gccgtgattg	14160
ccgacgccc	ccggtacaac	cccggaccgg	tgcctagcgc	cgcctccacc	gtggattccg	14220
tgatcgacag	cgtggtagct	ggcgtctggg	cctatgctcg	ccgcaagagg	cggtgcac	14280
ggagacgtcg	ccccaccgcc	gccatgctgg	cagccagggc	cgtgctgagg	cgggcccggg	14340
gggtaggcag	aagggtctatg	cgcgcgcgtg	ccgccaacgc	cgcgcgcggg	agggcccgc	14400
gacaggctgc	ccgccaggct	gctgccgcca	tgcctagcat	ggccagacc	aggagaggga	14460
acgtgtactg	ggtgcgcgat	tctgtgacgg	gagtcaggat	gccggtgcgc	agcgcacctc	14520
cccgaagtta	gaagatccaa	gctgcgaaga	cggcgggtact	gagtcctccct	gttgttatca	14580
gccaacatg	agcaagcgca	agtttaaaaga	agaactgctg	cagacgctgg	tgcctgagat	14640
ctatggccct	ccggacgtga	agcctgacat	taagccccgc	gatatcaagc	gtgttaaaaa	14700
gcgggaaaag	aaagagggaac	tgcggtgggt	agacgatggc	ggagtgggaat	ttattaggag	14760
tttcgccccg	cgacgcaggg	ttcaatggaa	agggcggcgg	gtacaacgcg	ttttgaggcc	14820
gggcaccgcg	gtagttttta	ccccgggaga	gcggtcggcc	gttaggggtt	tcaaaaggca	14880
gtacgacgag	gtgtacggcg	acgaggacat	attggaacag	gcggtcaac	agatcggaga	14940
atgtgcctac	ggaaagcggt	cgcgtcgcga	agacctggcc	atcgcttttag	acagcggcaa	15000
ccccacgccc	agcctcaaac	ctgtgacgct	gcagcagggtg	ctccccgtga	gcgccagcac	15060
ggacagcaag	aggggaataa	aaagagaaat	ggaagatctg	cagcccacca	tccagctcat	15120
ggctccctaaa	cggcagaggc	tggaagaggt	cctggagaaa	atgaaagtgg	acccaagcat	15180
agagccggac	gtcaaagtca	ggccgatcaa	agaagtggcc	cctggtctcg	gggtgcagac	15240
gggtgatata	cagatccccg	tcacgtcagc	ttcgaccgcc	gtggaagcca	tggaaacgca	15300
aacggaaaacc	cctgccgcga	tcggtaccag	ggaagtggcg	ttgcaaaccg	acccctggta	15360
cgaatacgcc	gcccctcggc	gtcagaggcg	acccgctcgt	tacggccccg	ccaacgccat	15420
catgccagaa	tatgcgctgc	atcgcgtctat	cctgcccacc	cccggctacc	ggggagtgc	15480
gtatcgcccc	tcaggaaacc	gcccgcgaac	ccgtcgccgc	cgcgcgtccc	gtcgtgctct	15540
ggcccccggtg	tcggtgcgcc	gcgtaacacg	ccggggaaag	acagttacca	ttcccaacc	15600
gcgctaccac	cctagcatcc	tttaatgact	ctgccgtttt	gcagatggct	ctgacttgcc	15660

gcgtgcgcct tcccgttccg cactatcgag gaagatctcg tcgtaggaga ggcattggcgg 15720
 gtagtggctg ccggcgggct ttgctgcaggc gcatgaaagg cggaaatttta cccgctctga 15780
 taccataat cgcgcgcgc atcggtgcca taccggcggt cgcttcagtg gccttgcaag 15840
 cagctcgtaa taaataaacg aaggcttttg cacttatgtc ctggtcctga ctattttatg 15900
 cagaaagagc atggaagaca tcaattttac gtcgctggct ccgcggcacg gctcgcggcc 15960
 gctcatgggc acctggaacg acatcggcac cagtcagctc aacgggggcg ctttcaattg 16020
 ggggagcctt tggagcggca ttaaaaactt tggctccacg attaaatcct acggcagcaa 16080
 agcctggaac agtagtgctg gtcagatgct ccgagataaa ctgaaggaca ccaacttcca 16140
 agaaaaagtg gtcaatgggg tggtagaccg catccacggc gcggtagatc tcgccaacca 16200
 agcggtagcag aaagagattg acaggcgctt ggaaagctcg cgggtgccgc cgcagagagg 16260
 ggatgaggtg gaggtcgagg aagtagaagt agaggaaaag ctgccccgcg tggagaaagt 16320
 tcccgtgctg cctccgagac cgcagaagcg acccaggcca gaactagaag aaactctggt 16380
 gacggagagc aaggagcctc cctcgtacga gcaagccttg aaagagggcg cctctccacc 16440
 ctaccaatg acaaaaccga tcgctgcctat ggctcggccg gtgtacggga aggactacaa 16500
 gcctgtcacg ctagagctcc cccgcgcgc accgcgcgc cccacgcgc cgacegttcc 16560
 cccccccctg ccggctccgt cggcgggacc cgtgtccgca cccgtcgcg tgcctctgcc 16620
 agccgcccgc ccagtggccg tggcactgc cagaaacccc agaggccaga gaggagccaa 16680
 ctggcaaagc acgctgaaca gcatcgtggg cctgggagtg aaaagcctga aacgcgcgcg 16740
 ttgctattat taaagtgtg gctaaaaaat ttcccgttgt atacgcctcc tatgttaccg 16800
 ccagagacgc gtgactgtcg ccgcgagcgc cgctttcaag atggccaccc catcgatgat 16860
 gccgcagtgg tcttacatgc acatcgccgg gcaggacgcc tcggagtacc tgagccccgg 16920
 tctcgtgcag ttgcgccgcg ccaccgacac ctacttcagc ttgggaaaca agtttagaaa 16980
 cccaccgtg gccccaccc acgatgtaac cacggaccgc tcgcaaaggc tgaccctgcg 17040
 ttttgtgccc gtagaccggg aggacaccgc gtactcttac aaagtgcgct acacgctggc 17100
 cgtaggggac aaccgagtgc tggacatggc cagcacctac tttgacatcc ggggagtgct 17160
 ggatcgcggt cccagtttta agccctactc gggtagcgcg tacaattccc tggctcccaa 17220
 gggcgctccc aaccctgcag aatggacgaa ttcagacagc aaagttaaag tgagggcaca 17280
 ggcgcctttt gttagctcgt atgggtgtac agcgattaca aaagagggtg ttcaggtggg 17340

agtaacctta	acagactccg	gatcaacacc	acagtatgca	gataaaacgt	atcagcctga	17400
gccgcaaatt	ggagaactac	agtggaacag	cgatgttggg	accgatgaca	aaatagcagg	17460
aagagtgcta	aagaaaacaa	cgcccatgtt	cccttggttac	ggctcatatg	ccaggcccac	17520
taatgaaaaa	ggaggacagg	caacacccgtc	cgctagtcaa	gacgtgcaaa	atccccgaatt	17580
acaatttttt	gcctctacta	atgtcgccaa	tacacaaaaa	gcagttctat	atgcggagga	17640
cgtgtcaatt	gaagcgccag	acactcactt	ggtgttcaaa	ccaacagtca	ctgaaggcat	17700
tacaagttca	gaggctctac	tgacccaaca	agctgctccc	aaccgtccaa	actacatagc	17760
ctttagagat	aattttattg	gtctcatgta	ctacaatagc	acaggtaaca	tgggagtact	17820
ggcaggccag	gcttctcagc	taaatgcagt	tgttgacctg	caagacagaa	atactgagct	17880
gtcctacca	ctcatgttgg	acgccctcgg	agaccgcagt	cggttactttt	ctatgtggaa	17940
ccaagctgtg	gatagttacg	atcctgatgt	aagaatcata	gaaaaccatg	gcgtagaaga	18000
tgaattgcct	aattattgct	ttcctttggg	aggcatggca	gtaaccgaca	cctactcgcc	18060
tataaagggt	aatggaggag	gcaatggatg	ggaagccaat	aacggcggtt	tcaccgaaag	18120
aggagtggaa	ataggttcag	ggaacatggt	tgccatggag	attaacctgc	aagccaacct	18180
atggcgtagc	tttctgtact	ccaatattgg	gctgtacctg	ccagactctc	tcaaaatcac	18240
tcctgacaac	atcacactcc	cagagaacaa	aaacacctat	cagtatatga	acggtcgcgt	18300
gacgccaccc	gggctgggtg	acacctacgt	taacgtgggc	gcgcgctggg	cccccgatgt	18360
catggacagt	attaaccttt	ttaatcacca	ccgcaacgcc	ggactccgct	accgttccat	18420
gctcctggga	aacggacgct	acgtgccctt	ccacatccag	gtgccccaga	aattctttgc	18480
aattaaaaac	ctgctgctgc	tccccggttc	ctacacctac	gagtggaaact	tccgcaagga	18540
cgtgaacatg	atcttgca	gctcgctggg	caatgacctg	cgagtggacg	gggccagcat	18600
ccgcttcgac	agcatcaacc	tgtacgccaa	ctttttcccc	atggcccaca	acacggcctc	18660
caccctggaa	gccatgctgc	gcaacgacac	caacgaccaa	tctttcaacg	actacctgtg	18720
cgcgccaac	atgctgtacc	ccatccccgc	caacgccacc	agcgtgcca	tctccattcc	18780
ctctcgcaac	tgggcagcct	tcaggggctg	gagtttcacc	cgcctcaaaa	ccaaggagac	18840
cccctcgctg	ggctccgggt	tcgaccccta	cttcgtctac	tccggctcca	tcccctacct	18900
ggacggcacc	ttctacctca	accatacttt	caaaaagggtg	tcaatcatgt	tcgactcctc	18960
cgtcagctgg	cccggcaacg	accgtctgct	gacgccaac	gagttcgaaa	tcaagcggtt	19020
ggtggacggg	gaagggtaca	acgtgggtca	gagcaacatg	accaaggact	ggttctctgat	19080

tcagatgctc agccactaca acatcggtta ccagggttcc tacgtgcccg aaaattacaa 19140
 ggaccgcatg tactctttct tcagaaactt ccaacccatg agccgccaaa ttgtagattc 19200
 aacggcttac actaattatc aggatgtgaa actgccatac cagcataaca actcaggggt 19260
 cgtgggttac atgggaccca ccattgcgaga ggggcaggcc taccggcca actatcccta 19320
 tccccgtatt ggggccaccg ccgtgccag cctcacgcag aaaaagtcc tctgcgaccg 19380
 ggtgatgtgg aggatccct tctctagcaa cttcatgtct atgggtccc tcaccgacct 19440
 ggggcagaac atgctgtacg ccaactccgc tcacgccttg gatatgacct ttgagggtgga 19500
 tcccatggat gagcccacgc ttctctatgt tctgtttgaa gtcttcgacg tgggtgcgcat 19560
 ccaccagccg caccgcggcg tcacgaggc cgtctacctg cgcacacctt tctctgccgg 19620
 taacgccacc acctaaagaa gccgatgggc tccagcgaac aggagctgca ggccattgtt 19680
 cgcgacctgg gctgcgggccc ctactttttg ggacacctcg acaagcgttt tcccggcttc 19740
 atgtccccc acaagccggc ctgtgccatc gtaaacacgg ccggacggga gaccgggggg 19800
 gtccactggc tcgccttcgc ctggaaccgc cgtaacgcga cctgctacct gttcgacct 19860
 tttggtttct ccgacgaaag gctgaagcag atctaccagt tcgagtacga ggggtctctc 19920
 aagcgcagcg ctctggcctc cagcccccag cactgcgtca ccctggaaaa gtccacccaa 19980
 acggtccagg gggccctctc ggcgcctgc gggctcttct gttgcatgtt tttgcacgcc 20040
 ttctgtcact ggccctcacac ccccatggat cacaacccca ccatggatct gctcaccgga 20100
 gtgcccacaa gcatgtttca cagccccag gtgcgccca ccctgcgccc taaccaggaa 20160
 cacctgtatc gctttctggg gaaactctt gcctattttt gccgccaccg gcagcgcac 20220
 gaacgggcca cggccttcga aagcatgagc caaagagtgt aatcaataaa aaacattttt 20280
 atttgacatg atacgcgctt ctggcgtttt attaaaaatc gaagggttcg agggaggggt 20340
 cctcgtgccc gctggggagg gacacgttgc gatactggaa acgggcgctc caacgaaact 20400
 cggggatcac cagccgcggc aggggcacgt cttctaggtt ctgcttccaa aactgccgca 20460
 ccagctgcag ggctcccatg acgtcgggcg ccgatatctt gaagtcgcag ttaggggccg 20520
 agctcccgcg gctgttgcgg aacacggggt tggcacactg gaacaccagc acgcccgggt 20580
 tgtggatact ggccagggcc gtgggtcgg tcacctccga cgcattcaga tctcggcgt 20640
 tgctcagggc aaacggggtc agcttgacac tctgccccc aatctggggt actaggctgc 20700
 gcttgttgag gcagtcgcag cgcagaggga tcaggatgcg tcgctgcccg cggttgcata 20760

tagggtaact	cgccgccagg	aactcctcca	tttgacggaa	ggccatctgg	gctttgccgc	20820
cctcgggtgta	gaatagcccg	caggacttgc	tagagaatac	gttatgaccg	cagttgacgt	20880
cctccgcgca	gcagcgggcg	tcttcgttct	tcagctgaac	cacgttgccg	ccccaacggt	20940
tctggaccac	cttggctcta	gtgggggtgct	ccttcagcgc	ccgctgtccg	ttctcgttgg	21000
ttacatccat	ttccaacacg	tgctccttgc	agaccatctc	cactccgtgg	aagcaaaaaca	21060
ggacgccctc	ctgctgggta	ctgcgatgct	cccatacggc	gcatccggtg	ggctcccagc	21120
tcttggtgtt	tacccccgcg	taggcttcca	tgtaagccat	aaggaatctg	cccatcagct	21180
cgggtgaagg	cttctgggtg	gtgaagggtta	gcggcaggcc	gcggtgctcc	tcgttcaacc	21240
aagtttgaca	gatcttgccg	tacaccgctc	cctggtcggg	cagaaactta	aaagccgctc	21300
tgctgtcggt	gtctacgtgg	aactttctcca	ttaacatcat	catggtttcc	atacccttct	21360
cccacgctgt	caccagtgg	ttgctgtcgg	ggttcttcac	caacacggcg	gtagaggggc	21420
cctcgcggc	cccgacgtcc	ttcatgggtca	ttctttgaaa	ctccacggag	ccgtccgcgc	21480
gacgtactct	gcgcaccgga	gggtagctga	agcccacctc	caccacgggtg	ccttcgccct	21540
cgtgtcgga	gacaatctcc	ggggatggcg	gcggcgccgg	tgctgccttg	cgagccttct	21600
tcttgggagg	gagctgaggc	gcctcctgct	cgcgctcggg	gctcatctcc	cgcaagtagg	21660
gggtaatgga	gctgcctgct	tggttctgac	ggttggccat	tgtatcctag	gcagaaagac	21720
atggagctta	tgcgcgagga	aactttaacc	gccccgtccc	ccgtcagcga	cgaagatgtc	21780
atcgtcgaac	aggaccggg	ctaagttacg	ccgcccagg	atctggagg	gcctgaccgg	21840
cgcgacgcta	gtgagcggca	ggaaaatgag	aaagaggagg	cctgctacct	cctggaaggc	21900
gacgttttgc	taaagcattt	cgccaggcag	agcaccatag	ttaaggaggc	cttgcaagac	21960
cgtccgagg	tgcccttgga	cgtcgccgcg	ctctcccagg	cctacgaggc	gaaccttttc	22020
tcgcctcgag	tgccctcgaa	gagacagccc	aacggcacct	gcgagcccaa	cccgcgactc	22080
aacttctacc	ccgtgttcgc	cgtaccagag	gcgctggcca	cctatcacat	ttttttcaaa	22140
aaccaacgca	tccccctatc	gtgcccggcc	aaccgcaccg	cggccgatag	gaatctcagg	22200
cttaaaaacg	gagccaacat	acctgatatc	acgtcgttgg	aggaagtgcc	caagattttc	22260
gaggggtctgg	gtcgagatga	gaagcgggcg	gcgaacgctc	tcgagaaaga	acagaaagag	22320
agtcagaacg	tgctggtgga	gctggagggg	gacaacgcgc	gtctggccgt	cctcaaacgc	22380
tgcatagaag	tctcccactt	cgcctacccc	gccctcaact	tgccacccaa	agttatgaaa	22440
tcgggtcatgg	atcagctgct	catcaagaga	gctgagcccc	tggtatccga	ccaccccgag	22500

gcggaaaact cagaggacgg aaagcccgtc gtcagcgacg aggagctcga gcggaggctg 22560
 gaaaccaggg accccaaca gttgcaagag aggcgcaaga tgatgatggc ggccgtgctg 22620
 gtcaccgtgg agctggaatg cctgcaacgg tttttcagcg acgtggagac gctacgcaaa 22680
 atcgggggaat ccctgcacta caccttccgc cagggctacg tccgccaggc ctgcaagatc 22740
 tccaacgtgg agctcagcaa cctgggtctcc tacatgggca tccctccacga gaaccggctg 22800
 gggcagagcg tgctgcactg caccttgcaa ggcgaggcgc ggcgggacta cgtgcgagac 22860
 tgcattctacc tcttctctac cctcacctgg cagaccgcca tggggcgtctg gcagcagtgc 22920
 ttggaagaga gaaacctcaa agagctagac aaactcctct gccgccagcg gcgcgccttg 22980
 tgggccgggt tcagcgagcg cacggctcgc agcgtcttg cggacatcat cttcccggag 23040
 cgcctgatga aaaccttgca aaacggcctg ccggatttca tcagtcaaag cattttgcaa 23100
 aacttccgct cttttgtcct ggaacgctcc gggatcttgc ccgccatgag ctgcgcgcta 23160
 ccttctgact ttgtccccct ctctaccgc gagtgcctc cccactgtg gagccactgc 23220
 tacctcttcc aactggccaa ctttctggcc taccactccg acctcatgga agacgtaagc 23280
 ggagaggggt tactggagtg ccactgccgc tgcaacctgt gcacccccca cagatcgctg 23340
 gcctgcaaca ccgagctact cagcgaaacc caggctatag gtaccttca gatccagggg 23400
 ccccagcagc aagaggggtg ttcgggcttg aagctactc cggcgctgtg gacctcggtc 23460
 tacttacgca aattttagc cgaggactac cacgcccaca aaattcagtt ttacgaagac 23520
 caatctcgac caccgaaagc cccctcacg gcctgcgtca tcaccagag caagatcctg 23580
 gcccaattgc aatccatcaa ccaagcgcg cgcgatttcc ttttgaaaaa gggtcggggg 23640
 gtgtacctgg acccccagac cggcgaggaa ctcaaccgt ccacactctc cgtcgaagca 23700
 gccccccga gacatgccgc ccaagggaac cgccaagcag ctgatcgctc ggcagagagc 23760
 gaagaagcaa gagctgctcc agcagcaggt ggaggacgag gaagagatgt gggacagcca 23820
 ggcagaggag gtgtcagagg acgaggagga gatggaaagc tgggacagcc tagacgagga 23880
 ggaggacgag ctttcagagg aagaggcgac cgaagaaaaa ccacctgcat ccagcgcgcc 23940
 ttctctgagc cgacagccga agccccggcc cccgacgcc ccggccggct cactcaaagc 24000
 cagccgtagg tgggacgcca ccgaatctcc agcggcagcg gcaacggcag cgggtaaggc 24060
 caaacgcgag cggcgggggg attgctcctg gcgggccac aaaagcagta ttgtgaactg 24120
 cttgcaacac tgcgggggaa acatctctt tgcccgacgc tacctcctct tccatcacgg 24180

tgtggccttc cctcgcaacg ttctctatta ttaccgtcat ctctacagcc cctacgaaac 24240
 gctcggagaa aaaagctaag gcttcctccg ccgcgaggaa aaactccgcc gccgctgccg 24300
 ccgccaagga tccaccggcc accgaagagc tgagaaagcg catctttccc actctgtatg 24360
 ctatctttca gcaaagccgc gggcagcacc ctacgcgcga actgaaaata aaaaaccgct 24420
 ccttccgctc gctcaccgcg agctgtctgt accacaagag agaagaccag ctgcagcgca 24480
 ccctggacga cgccgaagca ctgttcagca aatactgtc agcgtctctt aaagactaaa 24540
 agaccgcgcg tttttccccc tcggccgcca aaaccacgt catcgccagc atgagcaagg 24600
 agattcccac cccctacatg tggagctatc agccccagat gggcctggcc gccggggccg 24660
 cccaggacta ctccagcaag atgaactggc tcagcgccgg cccccacatg atctcacgag 24720
 ttaacggcat ccgagccac cgaaaccaga ttctcttaga acaggcggca atcaccgcca 24780
 cccccggcg ccaactcaac ccgcctagtt ggcccgccgc ccaggtgtat caggaaaatc 24840
 cccgcccgcg cacagtcctc ctgdcacgcg acgcggaggc cgaagtcctc atgactaact 24900
 ctggggtaca attagcgggc ggggtccaggt acgccaggta cagaggtcgg gccgctcctt 24960
 actctcccg gagtataaag aggggtgatc ttcgaggccg aggtatccag ctcaacgacg 25020
 agacggtgag ctctcaacc ggtctcagac ctgacggagt cttccagctc ggaggagcgg 25080
 gccgctcttc cttcaccact cgccaggcct acctgaccct gcagagctct tctcgcagc 25140
 cgcgctccgg gggaatcggc actctccagt tcgtggaaga gtctgttccc tctgtctact 25200
 tcaaccctt ctccggctcg cctggacgct acccgacgc cttcattccc aactttgacg 25260
 cagtgagtga atccgtggac ggctacgact gatgacagat ggtgcggccg tgagagctcg 25320
 gctgcgacat ctgcactact gccgtcagcc tcgtgctac gctcgggagg cgatcgtctt 25380
 cagctacttt gagctgccg acgagcacc tcagggtccg gctcacgggt tgaaactcga 25440
 gatcgagaac gcgctcgagt ctgcctcat cgacacctc accgcccgc ctctcctggt 25500
 agaaatcaa cgggggatca ctaccatcac cctgttctgc atctgcccc cgcccggtt 25560
 acatgaagat ctgtgtgtgc atctttgcgc tcagtttaat aaaaactgaa ctttttgccg 25620
 caccttcaac gccatctgtg atttctacaa caaaaagtgc ttctggcaaa ggtacacaaa 25680
 ctgtatttta ttctaattct acctcatcta tcgtgctgaa ctgcgcctgc actaacgaac 25740
 ttatccagtg gattgcaaac ggtagtgtgt gcaagtactt ttgggggaac gatatagtta 25800
 gtagaaataa cagcctttgc gagcactgca actcctccac actaatcctt tatcccccac 25860
 ttgttactgg atggtatatg tgcgttggt ccggtttaa tctagtgtg ttccataagt 25920

ggtttctaca aaaagagacc cttcccaaca attctgtttc ttttttcgcc ctatcctact 25980
 gctgttctcc ctctggttac tctttcaaac ctctaattgg tatttttagct ttgatactca 26040
 taatctttat taactttata ataattaaca acttacagta aacatgcttg ttctactgct 26100
 cgccacatct ttcgctctct ctcacgccag aacaagtatt gttggcgcag gttacaatgc 26160
 aactcttcaa tctgcttaca tgccagattc cgaccagata ccccatatta cgtgggtactt 26220
 acaaacctcc aaacctaatt cttcatttta tgaaggaaac aaactctgcg atgactccga 26280
 caacagaacg cacacatttc cccacccttc actacaattc gaatgcgtaa acaaaagctt 26340
 gaagctttac aacttaaagc cttcagattc tggcttgtag catgctgtag ttgaaaaaag 26400
 taatttagaa gtccacagtg attacattga attgacgggt gtggacctgc caccctccaa 26460
 atgtgagggt tctctctctt accttgaagt tcaaggcgtg gatgcctact gcctcataca 26520
 cattaactgc agcaactcta aatattccagc tagaatttac tataatggac aggaaagtaa 26580
 tcttttttat tatttaacaa caagcgttg taacggtaaa cagttacctg actattttac 26640
 tgcgtgtgtt gaattttcca cctacagaga aacgtatgcc aagcggcctt acaatttctc 26700
 ataccggtt aacgacctt gcaatgaaat acaagcgtc gaaactggaa ctgattttac 26760
 tccaattttc attgctgcca ttgttgtaag ctttaattacc attattgtca gcttagcatt 26820
 ttactgcttt tacaagcca aaaaccctaa gtttgaaaaa cttaaactaa aacctgtcat 26880
 tcaacaagtg tgattttgtt ttccagcatg gtagctgcat ttctacttct cctctgtcta 26940
 cccatcattt tcgtctcttc aactttcgcc gcagtttccc acctggaacc agagtgccta 27000
 ccgccttttg acgtgtatct gattctcacc tttgtttgtt gtatatccat ttgcagtata 27060
 gcctgctttt ttataacaat ctttcaagcc gccgactatt ttacgtgcg aattgcttac 27120
 tttagacacc atcctgaata cagaaatcaa aacgtgcct ccttactttg tttggcatga 27180
 ttaagttatt gctgatactt aattatttac ccctaataca ctgtaattgt ccattcacca 27240
 aaccctgggc attctacacc tggtatgata aaatccccga cactcctgtt gcttggcttt 27300
 acgcagccac cgccgctttg gtatttatat ctacttgcct tggagtaaaa ttgtatttta 27360
 ttttacacac tgggtggcta catcccagag aagatttacc tagatatcct cttgtaaacg 27420
 cttttcaatt acagcctctg cctcctcctg atcttcttcc tcgagctccc tctattgtga 27480
 gctactttca actcaccggt ggagatgact gactctcagg acattaatat tagtgtggaa 27540
 agaatagctg ctcagcgtca gcgagaaacg cgagtgttg aatacctgga actacagcaa 27600

cttaaagagt ccactggtg tgagaaagga gtgctgtgcc atgttaagca ggcagccctt 27660
 tcttacgatg tcagcgttca gggacatgaa ctgtcttaca ctttgccttt gcagaaacaa 27720
 accttctgca ccatgatggg ctctacctcc atcacaatca cccaacaagc cgggcctgta 27780
 gagggggcta tctctgtca ctgtcacgca cctgattgca tgtccaaact aatcaaaact 27840
 ctctgtgctt taggtgatat ttttaagggtg taaatcaata ataaacttac cttaaatttg 27900
 acaacaaatt tctggtgaca tcattcagca gcaccacttt accctcttcc cagctctcgt 27960
 atgggatgcg atagtgggtg gcaaacttcc tccaaaccct aaaagaaata ttggtatcca 28020
 cttccttgtc ctcaccaca attttcatct tttcatagat gaaaagaacc agagttgatg 28080
 aagacttcaa ccccgctctac ccttatgaca ccacaaccac tctgcagtt ccctttatat 28140
 cccccctt tgtaaacagc gatggtcttc aggaaaaccc ccaggtgtt ttaagtctgc 28200
 gaatagctaa acccctatat ttcgacatgg agagaaaact agccctttca cttggaagag 28260
 ggttgacaat taccgccgcc ggaqaattag aaagtacgca gagcgtacaa accaaccac 28320
 cgttgataat taccaacaac aacacactga ccctacgtca ttctcccccc ttaaacctaa 28380
 ctgacaatag cttagtcta ggctactcga gtctctccg cgtcacagac aacaaactta 28440
 catttaactt cacatcacca ctccgttatg aaaaataaaa ccttactttt aactatacag 28500
 agcctcttaa acttataaat aacagccttg ccattgacat caattcctca aaaggcctta 28560
 gtagcgtcgg aggctcacta gctgtaaacc tgagttcaga cttaaagttt gacagcaacg 28620
 gatccatagc ttttggcata caaacctgt ggaccgctcc gacctcgact ggcaactgca 28680
 cgtctacag cgagggcgat tccctactta gtctctgtt aaccaaatgc ggagctcacg 28740
 tcttaggaag tgtaagttta accgggttaa caggaaccat aaccctaatg actgatattt 28800
 ctgtcaccat tcaatttaca tttgacaaca atggtaagct actaagctct ccacttataa 28860
 acaacgcctt tagtattcga cagaatgaca gtacggcctc aaaccctacc tacaacgccc 28920
 tggcgtttat gcctaacagt accatatatg caagaggggg aggtggtgaa ccacgaaaca 28980
 actactacgt ccaaacgtat cttaggggaa atgttcaaaa accaatcatt cttactgtaa 29040
 cctacaactc agtcgccaca ggatattcct tatcttttaa gtggactgct cttgcacgtg 29100
 aaaagtttgc aaccccaaca acctcgtttt gctacattac agaacaataa aaccgtgtac 29160
 cccaccgttt cgtttttttc agatgaaacg ggcgagagtt gatgaagact tcaaccagat 29220
 gtacccttat gacccccac atgctcctgt tatgcccttc attactccac cttttacctc 29280
 ctcggatggg ttgcaggaaa aaccacttgg agtggttaagt ttaactaca gagatcccat 29340

tactacgcaa aatgagtctc ttacaattaa actaggaaac ggcctcactc tagacaacca	29400
gggacaacta acatcaaccg ctggcgaagt agaacctcca ctactaacg ctaacaacaa	29460
acttgactg gtctatagcg atcctttagc agtaaagcgc aacagcctaa ccttatcgca	29520
caccgctccc cttgttattg ctgataactc tttagcattg caagtttcag agcctatttt	29580
tataaatgac aaggacaaac tagccctgca aacagccgcg ccccttgtaa ctaacgctgg	29640
cacccttcgc ttacaaagcg ccgccccttt aggcattgca gaccaaacc taaaactcct	29700
gtttaccaac cctttgtact tgcagaataa cttttctacg ttagccattg aacgaccctt	29760
tgccattacc aatactggaa agctggctct acagctctcc ccaccgctac aaacagcaga	29820
cacaggcttg actttgcaaa ccaacgtgcc attaactgta agcaacggga ccctaggctt	29880
agccataaag cgcccactta ttattcagga caacaacttg tttttggact tcagagctcc	29940
cctgcgtctt ttcaacagcg acccagtact agggcttaac ttttacaccc ctcttgcggt	30000
acgcatgag gcgctcactg ttaacacagg ccgcggcctc acagtgagtt acgatgggtt	30060
aattttaaat cttggaagg atcttcgctt tgacaacaac accgtttctg tcgctcttag	30120
tgctgctttg cctttacaat acactgatca gcttcgcctt aacgtgggcg ctgggctgcg	30180
ttacaatcca gtgagtaaga aattggacgt gaaccccaat caaaacaagg gtttaacctg	30240
ggaaaatgac tacctcattg taaagctagg aaatggatta ggttttgatg gcgatggaaa	30300
catagctggt tctcctcaag ttacatcgcc tgacacctta tggaccactg ccgaccctc	30360
ccccaatgt tccatctaca ctgatttaga tgccaaaatg tggctctcgt tggtaaaaca	30420
aggggggtg gttcacggtt ctgttgcttt aaaagcattg aaaggaaccc tattgagtcc	30480
tacggaaagc gccattgtta ttatactaca ttttgacaat tatggagtgc gaattctcaa	30540
ttatcccact ttgggcactc aaggcacgtt gggaaataat gcaacttggg gttataggca	30600
gggagaatct gcagacacta atgtactcaa tgcactagca tttatgcca gttcaaaaag	30660
gtacccaaga gggcgtggaa gcgaagttca gaatcaaact gtgggctaca cttgtataca	30720
gggtgacttt tctatgcccg tacggtacca aatacagtac aactatggac caactggcta	30780
ctcctttaa tttatttgga gaactgttcc aagacaacca tttgacatcc catgctgttt	30840
tttctcttac attacggaag aataaaacaa ctttttcttt ttattttctt tttattttac	30900
acgcacagta aggttctctc cacccttcca tctcacagca tacaccagcc tctcccctt	30960
catggcagta aactgttggt agtcagtcgg gtatttggga gttaagatcc aaacagtctc	31020

ttggtgatg aaacatggat ccgtgatgga cacaaatccc tgggacaggt tctccaacgt 31080
 ttcggtaaaa aactgcatgc cgccctacaa aacaaacagg ttcaggctct ccacgggtta 31140
 tctccccgat caaactcaga cagagttaaag gtgcatgat gttccactaa accacgcagg 31200
 tggcgctgtc tgaacctctc ggtgcgactc ctgtgaggct ggtaagaagt tagattgtcc 31260
 agcagcctca cagcatggat catcagtcta cgagtgcgtc tggcgagca ggcacatctga 31320
 atctcactga gattccggca agaatcgac accatcacaa tcagggtgtt catgatccca 31380
 tagctgaaca cgctccagcc aaagctcatt cgctccaaca gcgccaccgc gtgtccgtcc 31440
 aaccttactt taacataaat caggtgtctg ccgctacaa acatgtacc cgcatacaga 31500
 acctccggg gcaaaccctt gttcaccacc tgctgtacc agggaaacct cacatattac 31560
 agggagccat agatagccat tttaaaccaa ttagctaaca ccgccccacc agctctacac 31620
 tgaagagaac cgggagagtt acaatgacag tgaataatcc atctctcata acccctaattg 31680
 gtctgatgga aatccagatc taacgtggca cagcagatac acactttcat atacattttc 31740
 atcacatgtt tttcccaggc cgttaaaata caatcccaat acacgggcca ctctgcagt 31800
 acaataaagc taatacaga tggtaactc ctacacctac taacattgtg catgttcata 31860
 ttttcacatt ctaagtaccg agagtctctc tctacaacag cactgccgcg gtcctcacia 31920
 ggtggttagct ggtgacgatt gtaaggagcc agtctgcagc gataccgtct gtcgcgttgc 31980
 atcgtagacc agggaccgac gcacttcctc gtacttgtag tagcagaacc acgtccgtg 32040
 ccagcacgtc tccaagtaac gccgggtcct gcgtcgctca cgctccctcc tcaacgcaaa 32100
 gtgcaaccac tcttgaatc cacacagatc cctctcggcc tccggggcga tgcacacctc 32160
 aaacctacag atgtctcggc acagttccaa acacgtagtg agggcgagtt ccaaccaaga 32220
 cagacagcct gatctatccc gacacactgg aggtggagga agacacgga gaggcattgt 32280
 attccaagcg attcaccaac gggtcgaaat gaagatcccg aagatgacaa cggtcgcctc 32340
 cggagccctg atggaattta acagccagat caaacattat gcgattttcc aggtatcaa 32400
 tcgcgccctc caaaagagcc tggaccgcga cttccacaaa caccagcaaa gcaaaagcgt 32460
 tattatcaaa ctcttcgac atcaagctgc aggactgtac aatgcccaag taattttcat 32520
 ttctccactc gcgaatgatg tcgcgcaaa tagtctgaag gttcatgccg tgcataataa 32580
 aaagctccga aaggcgccc tctatagcca tgcgtagaca caccatcatg actgcaagat 32640
 atcgggctcc tgagacacct gcagcagatt taacagaccc aggtcaggtt gctctccgcg 32700
 atcggaatc tccatccgca aagtcatttg caaataatta aatagatctg cgcgactaa 32760

atctgttaac	tccgcgctag	gaactaaatc	agggtgtggct	acgcagcaca	aaagtccag	32820
ggatggcgcc	aaactcacta	gaaccgctcc	cgagtagcaa	aactgatgaa	tgggagtaac	32880
acagtgtaaa	atgttcagcc	aaaaatcact	aagctgctcc	tttaaaaagt	ccagtacttc	32940
tatattcagt	tcgtgcaagt	actgaagcaa	ctgtgcgga	atatgcacag	caaaaaaat	33000
agggcggtc	agatacatgt	tgacctaaaa	taaaaagaat	cattaaacta	aagaagcctg	33060
gcgaacggtg	ggatatatga	cacgctccag	cagcaggcaa	gcaaccggct	gtccccggga	33120
accgcggtaa	aattcatccg	aatgattaaa	aagaacaaca	gagacttccc	accatgtact	33180
cggttggatc	tcctgagcac	agagcaatac	ccccctcaca	ttcatatccg	ctacagaaaa	33240
aaaacgtccc	agataccag	cgggaatatc	caacgacagc	tgcaaagaca	gcaaaaacaat	33300
ccctctggga	gcaatcacia	aatcctccgg	tgaaaaaagc	acatacatat	tagaataacc	33360
ctgttgctgg	ggcaaaaagg	cccgtcgtcc	cagcaaatac	acataaatat	gttcatcagc	33420
cattgccccg	tcttaccgcg	taaaagagcca	cgaaaaaatc	gagctaaaat	ccaccacaaca	33480
gcctatagct	atatatacac	tccaccaaat	gacgctaata	ccgcaccacc	cacgacaaaa	33540
gttcacccac	accacaaaaa	cccgcgaaaa	tccagcgccg	tcagcacttc	cgcaatttca	33600
gtctcacaac	gtcacttcg	cgcgcctttt	cactttccca	cacacgcctt	tcgcccgcgc	33660
gccctcgcgc	caccccggt	cacccacagt	caccgcacgt	caccccggtc	ccgcctcgtc	33720
cctccccgct	cattatcata	ttggcacgtt	tccagaataa	ggtatattat	tgatgcagca	33780
aaacaatccc	tctgggagca	atcacaaaat	cctccggtga	aaaaagcaca	tacatattag	33840
aataaccctg	ttgctggggc	aaaaaggccc	gtcgtcccag	caaatgcaca	taaatatggt	33900
catcagccat	tgccccgtct	taccgcgtaa	acagccacga	aaaaatcgag	ctaaaatcca	33960
cccaacagcc	tatagctata	tatacactcc	acccaatgac	gctaataccg	caccacccac	34020
gaccaaagtt	cacccacacc	cacaaaaccc	gcgaaaatcc	agcgccgtca	gcacttcgcg	34080
aatttcagtc	tcacaacgtc	acttcgcgc	gccttttcac	tttccacac	acgcccttcg	34140
cccgcgcgc	ctcgcgccac	cccgcgtcac	cccacgtcac	cgcacgtcac	cccggccccg	34200
cctcgtcctt	ccccgctcat	tatcatattg	gcacgtttcc	agaataaggt	atattattga	34260
tgca						34264

<210> 5
 <211> 31044
 <212> DNA

<213> simian adenovirus SV-25

<400> 5

catcatcaat aatatacctt attctggaaa cgtgccata tgataatgag cggggaggag	60
cgaggcgggg ccggggtgac gtgcggtgac gcggggtggc gcgagggcgg ggcgaagggc	120
gcgggtgtgt gtgtgggagg cgcttagttt ttacgtatgc ggaaggagggt ttataccgg	180
aagatgggta atttgggcgt atacttgtaa gttttgtgta atttggcgcg aaaactgggt	240
aatgaggaag ttgagggttaa tatgtacttt ttatgactgg gcggaatttc tgctgatcag	300
cagtgaactt tgggcgctga cggggagggt tcgctacgtg acagtaccac gagaaggctc	360
aaaggcccc tttattgtac tcttcagcgt ttctcgctggg tatttaaacg ctgtcagatc	420
atcaagaggc cactcttgag tgctggcgag aagagttttc tcctccgtgc tgccacgatg	480
aggctggtcc ccgagatgta cgggtgtttt agcgacgaga cggtgcgtaa ctcatatgac	540
ctgctgaatt cagacgcgct ggaaatttcc aattcgctg tgctttcgcc gccgtcactt	600
cacgacctgt ttgtgttttg gctcaacgct tagcaacgtg ttatataggg tcaagaagga	660
gcaggagacg cagtttgcta ggctgttggc cgatactcct ggagtttttg tggctctgga	720
tctaggccat cactctcttt tccaagagaa aattatcaaa aacttaactt ttacgtctcc	780
tggtcgacg gttgcttcg ctgcctttat tacctatatt ttggatcaat ggagcaacag	840
cgacagccac ctgtcgctgg agtacatgct ggattacatg tcgatggcgc tgtggagggc	900
catgctgcgg agggagggtt gcatttactt gcgggcgcag cctccgcggc tggaccgagt	960
ggaggaggag gacgagccgg gggagaccga gaacctgagg gccgggctgg accctccaac	1020
ggaggactag gtgctgagga tgatcccgaa gaggggacta gtggggctag gaagaagcaa	1080
aagactgagt ctgaacctcg aaactttttg aatgagttga ctgtgagttt gatgaatcgt	1140
cagcgtccgg agacaatttt ctgggtctgaa ttggaggagg aattcaggag gggggaactg	1200
aacctgctat acaagtatgg gtttgaacag ttaaaaactc actggttgga gccgtgggag	1260
gattttgaaa ccgccttgga cacttttgct aaagtggctc tcgggccgga taaggtttac	1320
actatccgcc gcactgttaa cataaagaag agtgtttatg ttataggcca tggagctctg	1380
gtgcagggtc aaaccgtcga ccgggtggcc tttagttgag gtatgcaaaa tctgggcccc	1440
gggggatag gcttaaattg tgtaacattt cacaatgtaa ggtttactgg tgaaagtttt	1500
aacggctctg tgtttgcaaa taacacacag ctgacgctcc acggcgttta cttttttaac	1560
tttaataaca catgtgtgga gtcgtggggc aggggtgtctt tgaggggctg ctgttttcac	1620

ggctgctgga aggcgggtggt ggggaagactt aaaagtgtaa catctgtaaa aaaatgcggtg	1680
tttgagcgggt gtgtgttggc ttttaactgtg gagggctgtg gacgcattag gaataatgcg	1740
gcgtctgaga atggatgttt tcttttgcta aaaggcacgg ctagtattaa gcataacatg	1800
atatgcggca gcggtctgta cccttcacag ctgttaactt gcgcggatgg aaactgtcag	1860
accttgcgca ccgtgcacat agcgtccac cagcgccgcg cctggccaac attcgagcac	1920
aatatgctta tgcgttgtgc cgtccacttg ggccttaggc gaggcgtgtt tgtgccttac	1980
cagtgtaaact ttagccatac caagatttta ctagaacctg ataccttctc tctgagtgtgt	2040
ttcaatgggg tgtttgacat gtcaatggaa ctgttttaaag tgataagata tgatgaatcc	2100
aagtctcggt gtcgcccattg tgaatgcgga gctaatactc tgaggttgta tccctgtaacc	2160
ctaaacgcta ccgaggagct gaggacggat caccacatgt tgtcctgcct gcgcaccgac	2220
tatgaatcca gcgacgagga gtgagggtgag ggcgggagcc acaaagggtg taaaggggcg	2280
tgaggggtgg gtgtgatgat tcaaaatgag cgggacgacg gacggcaacg cgtttgaggg	2340
tggagtgttc agcccttate tgacatctcg tcttccttcc tgggcaggag tgcgtcagaa	2400
tgtagtgggc tccaccgtgg acggacgacc ggtcgccctt gcaaattccg ccaccctcac	2460
ctatgccacc gtgggatcat cgttgacac tgcgcgggca gctgccgctt ctgctgccgc	2520
ttctactgct cgcggcatgg cggctgattt tggactgtat aaccaactgg cctactgcagc	2580
tgtggcgctc cggctctctg ttcaagaaga tgcctgaat gtgatcctga ctgcctgga	2640
gatcatgtca cgtcgcttgg acgaactggc tgcgcagata tcccaagcta accccgatac	2700
cacttcagaa tctaaaaata aagacaaaca aatatgttga aaagtaaaat ggctttattt	2760
gttttttttg gctcggtagg ctcggttcca cctgtctcgg tgcgttaagaa ctttgtgtat	2820
gttttccaaa acacggtaca gatgggcttg gatgttcaag tacatgggca tgaggccatc	2880
tttgggggtga agataggacc attgaagagc gtcattgctc ggggtggtgt tgtaaattac	2940
ccagtcttag cagggtttct gggcgtggaa ctggaagatg tcccttagga gtaggctgat	3000
ggccaagggc aggccttag tgtaggtgtt tacaaagcgg ttaagctggg agggatgcat	3060
gcgggggggag atgatatgca tcttggttg gatcttgagg ttagctatgt taccaccag	3120
gtctctgcgg gggttcatgt tatgaaggac caccagcacg gtgtagccgg tgcatttggg	3180
gaacttgtca tgcagtttgg aggggaaggc gtggaagaat ttagagaccc ccttggtggc	3240
ccctaggttt tccatgcact catccataat gatggcaatg ggaccctgg cggccgcttt	3300
ggcaaacacg ttttgggggt tggaaacatc atagttttgc tctagagtga gctcatcata	3360

ggccatctta acaaagcggg gtaggagggg gcccgactgg gggatgatag ttccatctgg	3420
gcctggggcg tagttaccct cacagatctg catctcccag gccttaattt ccgagggggg	3480
tatcatgtcc acctgggggg caataaagaa cacggtttct ggcgggggat tgatgagctg	3540
ggtggaaagc aagttacgca gcagttgaga tttgccacag ccggtggggc cgtagatgac	3600
cccgatgacg ggttgacgct ggtagttgag agaggaacag ctgccgtcgg ggcgcaggag	3660
gggggctacc tcattcatca tgcttctaac atgtttattt tcaactacta agttttgcaa	3720
gagcctctcc ccaccagggt ataagagttc ttccaggctg ttgaagtgtt tcagcggttt	3780
taggccgtcg gccatgggca tcttttcgag cgactgacga agcaagtaca gtcggtccca	3840
gagctcggtg acgtgctcta tggaatctcg atccagcaga cttcttggtt gcggggggtg	3900
ggtcgacttt cgctgtaggg caccagccgg tgggcgtcca gggccgcgag ggttctgtcc	3960
ttccagggtc tcagcgctccg ggtgagggtg gtctcgggtga cgggtgaagggt atgagccccg	4020
ggctggggcg ttgcgagggt gcgcttcagg ctcatcctgc tgggtgctgaa gcggacgtcg	4080
tctccctgtg agtcggccag atagcaacga agcatgaggt cgtagctgag ggactcggcc	4140
gcgtgtccct tggcgcgcag ctttcccttg gaaacgtgct gacatttggt gcagtgcaga	4200
cattggagggt cgtagagttt gggggccagg aagaccgact cgggcgagta ggcgtcggct	4260
ccgcactgag cgcagacggt ctgcactcc actagccacg tgagctcggg tttagcggga	4320
tcaaaaacca agttgcctcc attttttttg atgcgtttct taccttgctt ttccatgagt	4380
ttgtggcccc cttccgtgac aaaaaggctg tcggtgtctc cgtagacaga cttgaggggg	4440
cgatcttcca aagggtgttc gaggtcttcc gcgtacagga actgggacca ctccgagacg	4500
aaggctctgg tcagggttaa cacgaaggag gcaatctgcg aggggtatct gtcgttttca	4560
atgaggggggt ccaccttttc cagggtgtgc agacacagggt cgtcctctcc cgcgctccacg	4620
aagggtgattg gcttgtaagt gtaggtcacg tgatctgcac ccccaaagg ggtataaaag	4680
ggggcggtgc caccctctcc gtcactttct tccgcctcgc tgtggaccag agccagctgt	4740
tggggtgagt aggcctctc aaaagccggc atgatctcgg cgtcaagtt gtcagtttct	4800
acaaacgagg tggatttgat attcacgtgc cccgcggcga tgcttttgat ggtggagggg	4860
ttcatctgat cagaaaacac gatctttttg ttgtcaagtt tgggtggcgaa agaccgtag	4920
agggcgttgg aaagcaactt ggcgatggag cgcagggtct gatttttctc ccgatcggcc	4980
ctctccttgg cggcgatggt gagggtgcacg tactcccggt ccgcgcaccg ccactcgggg	5040

aacacggcgg tgcgctcgtc gggcaggatg cgcacgcgcc agccgcgatt gtgcaggggtg	5100
atgaggtcca cgctggtagc cacctccccg cggaggggct cgttggtcca acacaatcg	5160
cccccttttc tggagcagaa cggaggcagg ggatctagca agttggcggg cggggggtcg	5220
gcgtcgatgg tgaagatacc gggtagcagg atcttattaa aataatcgat ttcggtgtcc	5280
gtgtcttgca acgcgtcttc ccacttcttc accgccaggg ccctttcgta gggattcagg	5340
ggcgggtccc agggcatggg gtgggtcagg gccgaggcgt acatgccgca gatgtcatac	5400
acgtacaggg gtccctcaa caccgcgatg taagtgggt aacagcgccc ccgcgggatg	5460
ctggctcgca cgtagtcgta catctcgcg caggagacca tgaggccgtc tccaagtgg	5520
gtcttggtgg gtttttcggc ccggtagagg atctgtctga agatggcgtg ggagttggaa	5580
gagatgggtgg ggcgttgaa gacgttaaag ttggccccg gtagtccac ggagtcttg	5640
atgaactggg cgtaggattc ccggagttt tccaccaggg cggcggtcac cagcacgtcg	5700
agagcgcagt agtccaacgt ctgcggacc aggttgtagg ccgtctcttg tttttctcc	5760
cacagttcgc ggttgaggag gtattcctcg cggcttttc agtactcttc ggcgggaaat	5820
ctttttctgt ccgctcggta agaacctaac atgtaaaatt cgttaccgc tttgtatgga	5880
caacagcctt tttctaccg cagggcgtag gcttgagcgg cctttctgag agaggtgtgg	5940
gtgagggcga aggtgtccc caccatcact ttcaggtact gatgtttgaa gtccgtgtcg	6000
tcgcaggcgc cctgttccca cagcgtgaag tcggtgcgt ttttctgcct gggattgggg	6060
agggcggaagg tgacatcgtt aaagagtatt tccccggcg ggggcattgaa gttgcgagag	6120
atcctgaagg gcccgggcac gtccgagcgg ttggtgatga cctgcgccgc caggacgac	6180
tcgtcgaagc cgttgatgtt gtgaccacg atgtaaagt cgatgaagcg cggctgtccc	6240
ttgagggccg gcgcttttt caactcctcg taggtgagac agtcggcga ggagagaccc	6300
agctcagccc gggcccagtc ggagagttga ggattagccg caaggaagga gtcctataga	6360
tccaaggcca ggagagtttg caagcggtc cggaaactcg ggaactttt cccacggcc	6420
attttctcgg gtgtcactac gtaaaagggt ttggggcggt tgttccacac gtccatcgg	6480
agctctaggg ccagctcgca ggcttggcga acgagggct cctcgccaga gacgtgcatg	6540
accagcataa aggttaccaa ctgtttcccg aacgagccca tccatgtgta ggtttctacg	6600
tcgtaggtga caaagagccg ctgggtgcgc gcgtgggagc cgatcggaag gaagctgatc	6660
tcctgccacc agctggagga atgggtgtta atgtgtgga agtagaagtc ccgcggcg	6720
acagagcatt cgtgctgatg tttgtaaaag cgaccgcagt agtcgcagcg ctgcacgctc	6780

tgtatctcct gaacgagatg cgcttttcgc ccgcgcacca gaaaccggag ggggaagtgtg	6840
agacgggggg ctgggtgggc gacatccccct tcgccttggc ggtgggagtc tgcgtctgcg	6900
tcctccttct ctgggtggac gacgggtggg acgacgacgc cccgggtgcc gcaagtccag	6960
atctccgcca cggaggggtg caggcgctgc aggaggggac gcaçctgccc gctgtccagg	7020
gagtcgaggg aagtgcgct gaggtcggcg ggaagcgttt gcaagttcac tttcagaaga	7080
ccggtaagag cgtgagccag gtgcagatgg tacttgattt ccaggggggt gttggatgaa	7140
gcgtccacgg cgtagaggag tccgtgtccg ccggggggcca ccaccgtgcc ccgaggagggt	7200
tttatctcac tcgtcgaggg cgagçgccgg ggggtagagg cggctctgcg ccggggggca	7260
gcggaggcag aggcacgttt tcgtgaggat tcggcagcgg ttgatgacga gcccgagac	7320
tgctggcggtg ggcgacgacg cggcggttga ggtcctggat gtgccgtctc tgcgtgaaga	7380
ccaccggccc ccgggtcctg aacctaaaga gagttccaca gaatcaatgt ctgcatcgct	7440
aacggcgggc tgcctgagga tctcctgcac gtcgcccag ttgtcctgat aggcgatctc	7500
ggccatgaac tgttccactt ctccctcgcg gaggtcaccg tggcccgctc gctccacgggt	7560
ggcgcccagg tcgttgagga tgcggcgcat gagttgagag aaggcggtga ggccgttctc	7620
gttccacacg cggctgtaca ccacgtttcc gaaggagtgc cgcgctcgca tgaccacctg	7680
ggccacgttg agttccacgt ggcgggcgaa gacggcgtag tttctgaggc gctggaagag	7740
gtagttgagc gtggtggcga tgtgctcgca gacgaagaag tacataatcc agcgccgcag	7800
ggtcatctcg ttgatgtctc cgatggcttc gagacgctcc atggcctcgt agaagtcgac	7860
ggcgaagtgt aaaaattggg agttgcgggc ggccaccgtg agttcttctt gcaggaggcg	7920
gatgagatcg gcgaccgtgt cgcgcacctc ctgttcgaaa gcgccccgag gcgcctctgc	7980
ttcttctctc ggctcctcct ctccagggg ctcgggttcc tccggcagct ctgcgacggg	8040
gacggggcgg cgacgtcgtc gtctgaccgg caggçggtcc acgaagcgct cgatcatttc	8100
gccgcgccgg cgacgcagtg tctcggtgac ggcgcgctcc ttttcgagag gtcgcagttc	8160
gaagaçgccg ccgcgcagag cggcccgctg cagggaggggt aagtggttag ggccgtcggg	8220
cagggaacac gcgctgacga tgcattttat caattgctgc gtaggcactc cgtgcagggg	8280
tctgagaacg tcgaggtcga cgggatccga gaacttctct aggaaagcgt ctatccaatc	8340
gcaatcgcaa ggtaagctga gaacgggtggg tcgctggggg gcgttcgcgg gcagttggga	8400
ggtgatgctg ctgatgatgt aattaaagta ggcggtcttc aggcggcgga tgggtggcgag	8460

gaggaccacg tctttgggcc cggcctgttg aatgcgagc cgctcggcca tgcccaggc	8520
ctcgctctga cagcgacgca ggtctttgta gaagtcttgc atcagtctct ccaccggaac	8580
ctctgcttct cccctgtctg ccatgcgagt cgagccgaac ccccgaggg gctgcagcaa	8640
cgctagggtcg gccacgaccc ttccggccag caccggcctgt tgaatctgag tgaggggtggc	8700
ctggaagtcg tccaggtcca cgaagcgggtg ataggccccc gtgttgatgg tgtaggtgca	8760
gttggccatg acggaccagt tgacgacttg catgccgggt tgggtgatct ccgtgtactt	8820
gaggcgag tagggcctgg actcgaacac gtagtcgttg catgtgcgca ccagatactg	8880
gtagccgacc aggaagttag gaggcggtc tcggtacagg ggccagccaa cgggtggcggg	8940
ggcgccgggg gacaggctcg ccagcatgag gcggtggttag tggtagatgt agcgggagag	9000
ccagggtgat ccggccgagg tgggtgcggc cctggtgaat tcgcggacgc ggttccagat	9060
gttgcgaggg ggaccaaagc gctccatggt gggcacgctc tgccccgtga ggcgggcgca	9120
atcttgtagc ctctagatgg aaaaagaca gggcggtcat cgactccttt ccgtagcttg	9180
gggggtaaag tcgcaagggg gggcgcgcg ggaaccccgg ttcgagaccg gccggatccg	9240
ccgctccga tgcgcctggc cccgcatcca cgacgtccgc gccgagacc agccgcgacg	9300
ctccgcccc atacggaggg gagtcttttg gtgttttttc gtagatgcat ccggtgctgc	9360
ggcagatgag accccagacg cccactacca ccgccgtggc ggcagtaa acctgagcggag	9420
gcggtgacag ggaggaggaa gagctggctt tagacctgga agaggagag gggtggccc	9480
ggctgggagc gccatcccc gagagacacc ctagggttca gctcgtgagg gacgccaggc	9540
aggcttttgt gccgaagcag aacctgttta gggaccgcag cggtcaggag gcggaggaga	9600
tgcgcgattg caggtttcgg gcgggcagag agctcagggc gggcttcgat cgggagcggc	9660
tcctgagggc ggaggatttc gagccgacg agcgttctgg ggtgagccc gcccgcgctc	9720
acgtatcggc ggccaacctg gtgagcgcgt acgagcagac ggtgaacgag gagcgcaact	9780
tccaaaagag ctttaacaat cacgtgagga ccctgatcgc gaggagagg gtgacctcg	9840
ggctgatgca tctgtgggac ttcgtggagg cctacgtgca gaacccggct agcaaacc	9900
tgacggccca gctgttcctg atcgtgcagc acagccgca caacgagacg ttccgcgacg	9960
ccatgttgaa catcgaggag cccgagggtc gctggctctt ggatctgatt aacatcctgc	10020
agagcatcgt ggtgcaggag aggggcctga gtttagcgga caagggtggc gccattaact	10080
attcgatgca gagcctgggg aagttctacg ctgcgaagat ctacaagagc ccttacgtgc	10140
ccatagacaa ggaggagaag atagacagct ttacatgcg catggcgctg aagggtgctga	10200

cgctgagcga cgacctcggc gtgtaccgta acgacaagat ccacaaggcg gtgagcgcca 10260
 gccgcccggcg ggagctgagc gacagggagc tgatgcacag cctgcagagg gcgctggcgg 10320
 gcgcccggga cgaggagcgc gaggcttact tgcacatggg agccgatctg cagtggcgtc 10380
 ccagcgcgcg cgccttgag gggcggggtt atcccgcga ggaggatcgg gacgatttgg 10440
 aggaggcagg cgagtacgag gacgaagcct gaccgggcag gtgttgtttt agatgcagcg 10500
 gccggcggac gggaccaccg cggatcccgc acttttggca tccatgcaga gtcaaccttc 10560
 gggcgtgacc gcctccgatg actgggcggc ggccatggac cgcacatgg cgctgaccac 10620
 ccgcaacccc gaggctttta ggcagcaacc ccaggccaac cgtttttcgg ccattcttga 10680
 agcgggtgtg ccgtcgcgca ccaaccgac gcacgagaaa gtccctgacta tcgtgaacgc 10740
 cctggtagac agcaaggcca tccgcctga cgaggcgggc ttgatttaca acgctctttt 10800
 ggaacgcgtg gcgcgtaca acagcactaa cgtgcagacc aatctggacc gcctcaccac 10860
 cgacgtgaag gaggcgctgg cgcagaagga gcggtttctg agggacagta atctgggctc 10920
 tctgggtggca ctgaaegcct tcctgagctc acagccggcc aacgtgcccc gcgggcagga 10980
 ggattacgtg agcttcatca gcgctctgag actgctggtg tccgaggtgc cccagagcga 11040
 ggtgtaccag tctgggcccgg attacttttt ccagacgtcc cgacagggct tgcaaaccgt 11100
 gaacctgact caggccttta aaaacttgca aggcattgtg ggggtcaagg ccccggtggg 11160
 cgatcgcgcc actatctcca gtctgctgac cccaacact cgcctgctgc tgcctcttgat 11220
 cgcaccgttt accaacagta gcactatcag ccgtgactcg tacctgggtc atctcatcac 11280
 tctgtaccgc gaggccatcg gccaggctca gatcgacgag catacgtatc aggagattac 11340
 taacgtgagc cgtgccctgg gtcaggaaga taccggcagc ctggaagcca cgttgaactt 11400
 tttgctaacc aaccggaggc aaaaaatacc ctcccagttc acgttaagcg ccgaggagga 11460
 gaggattctg cgatacgtgc agcagtcctg gagcctgtac ttgatgcgcg agggcgccac 11520
 cgcttccacg gcttttagaca tgacggctcg gaacatggaa ccgtcctttt actccgcccc 11580
 ccggccgttc attaacgtc tgatggacta ctccatcgc gcggccgcca tgaacgggga 11640
 gtadttcacc aatgccatcc tgaatccgca ttggatgccc ccgtccggct tctacaccgg 11700
 ggagtttgac ctgccgaag ccgacgacgg ctttctgtgg gacgacgtgt ccgatagcat 11760
 tttcacgccg gctaatacgc gattccagaa gaaggagggc ggagacgagc tccccctctc 11820
 cagcgtggaa gcggcctcaa ggggagagag tccctttcca agtctgtctt ccgccagtag 11880

cggtcgggta acgcgtccac ggttgccggg ggagagcgac tacctgaacg accccttget 11940
 gcgaccggct agaaagaaaa attttcccaa taacgggggtg gaaagcttgg tggataaaat 12000
 gaatcgttgg aagacgtacg cccaggagca gcgggagtgaggacagtc agccgcggcc 12060
 gctgggtaccg ccgcattggc gtcgccagag agaagaccctg gacgactccg cagacgatag 12120
 tagcgtgttg gacctgggag ggagcggagc caacccttt gctcacttgc aaccaagggt 12180
 gcgctcgagt cgctgtatt aataaaaaag acgcggaaac ttaccagagc catggccaca 12240
 gcgtgtgtgc tttcttctc ttttcttcc tcggcgcggc agaattgagaa gagcgggtgag 12300
 agtcacgccg gcggcgatg agggcccgcc cccttcttac gaaagcgtga tgggatcagc 12360
 gaacgtgccg gccacgctgg aggcgcctta cgttctctcc agatacctgg gacctacgga 12420
 gggcagaaac agcatccgtt actccgagct ggcgcctctg tacgatacca ccaagggtgta 12480
 cctgggtggac aacaagtcgg cggacatcgc ctccctgaat taccaaaacg atcacagtaa 12540
 ctttctgact accgtggtgc agaacaatga cttcaccctg acggaggcgg gcacgcagac 12600
 cattaacttt gacgagcgtt cccgctgggg cggtcagctg aaaaccatcc tgcacaccaa 12660
 catgcccac atcaacgagt tcatgtccac caacaagttc agggctaagc tgatggtaga 12720
 aaaaagtaat gcggaaactc ggcagccccg atacgagtgg ttcgagttta ccattccaga 12780
 gggcaactat tccgaaacta tgactatcga tctcatgaat aacgcgatcg tggacaatta 12840
 cctgcaagtg gggagacaga acggggtgct ggaaagcgat atcggcgatga aattcgatac 12900
 cagaaacttc cgactggggt gggatcccgt gaccaagctg gtgatgccag gcgtgtacac 12960
 caacgaggct tttcaccctg acatcgtgct gctgccgggg tgcgggtgtg acttcaactca 13020
 gagccgtttg agtaacctgt taggaattag aaagcgccgc cccttccaag agggctttca 13080
 aatcatgtat gaggacctgg agggaggtaa tatacccgcc ttactggacg tgtcgaagta 13140
 cgaagctagc atacaacgcg ccaaagcgga gggtagagag attcggggag acacctttgc 13200
 ggtagctccc caggacctgg aaatagtgcc tttaactaaa gacagcaaag acagaagcta 13260
 caatattata aacaacacga cggacacct gtatcgagc tggtttctg cttacaacta 13320
 cggagacccc gagaaaggag tgagatcatg gaccatactc accaccacgg acgtgacctg 13380
 tggctcgcag caagtgtact ggtccctgcc ggatatgatg caagaccggg tcaccttccg 13440
 cccctccacc caagtcagca acttcccggg ggtgggcacc gagctgctgc ccgtccatgc 13500
 caagagcttc tacaacgagc aggcggtcta ctgcgaactt attcgccagt ccaccgcgct 13560
 taccacgtg ttcaatcgt ttcccagaaa ccagattctg gtgcgcctc ccgtcctac 13620

cattaccacc gtcagtgaac acgttcccgc cctcacagat cacggaaccc tgccgctgcg 13680
cagcagtatc agtggagttc agcgcgtgac catcaccgac gccagacgtc gaacctgccc 13740
ctacgtttac aaagcgcttg gcgtgggtggc tcctaaagtt ctttctagtc gcaccttcta 13800
aaaacatgtc catcctcatc tctcccata acaacaccgg ctggggactg ggctccggca 13860
agatgtacgg cggagccaaa aggcgctcca gtcagcacc agttcgagtt cggggccact 13920
tccgcgctcc ttggggagct tacaagcgag gactctcggg tcgaacggct gtagacgata 13980
ccatagatgc cgtgattgcc gacgcccggc ggtacaaccc cggaccggtc gctagcggcg 14040
cctccaccgt ggattccgtg atcgacagcg tggtagccgg cgtcggggcc tatgctcgcc 14100
gcaagaggcg gctgcatcgg agacgtcgcc ccaccgccc catgctggca gccaggggcg 14160
tgctgaggcg ggcccggagg gcaggcagaa gggctatgcg ccgcgctgcc gccaacggcg 14220
ccgcggggag ggcccgcga caggctgcc gccaggctgc cgtgccatc gctagcatgg 14280
ccagaccag gagagggaac gtgtactggg tgcgtgattc tgtgacggga gtccgagtcg 14340
cggtgcgag ccgacctccc cgaagttaga agatccaagc tgcaagacg gcggtactga 14400
gtctccctgt tgttatcagc ccaacatgag caagcgcaag ttaagaag aactgctgca 14460
gacgtgggt cctgagatct atggccctcc ggacgtgaag ccagacatta agccccgca 14520
tatcaagcgt gttaaaaagc gggaaaagaa agaggaaactc gcggtggtag acgatggcg 14580
agtgaattt attaggagtt tcgccccgcg acgcagggtt caatggaaag ggcggcggt 14640
acaacgcgtt ttgaggccgg gcaccgggt agttttacc ccgggagagc ggtcggccgt 14700
taggggtttc aaaaggcagt acgacgaggt gtacggcgac gaggacatat tggaacaggc 14760
ggctcaacag atcggagaat ttgctacgg aaagcgttcg cgtcggaag acctggccat 14820
cgcttagac agcggcaacc ccacggccag cctcaaaccc gtgacgctgc agcaggtgct 14880
tcccgtagc gccagcacgg acagcaagag ggggattaag agagaaatgg aagatctgca 14940
tcccaccatc caactcatgg tcctaaacg gcagaggctg gaagaggtcc tggagaagat 15000
gaaagtggac ccagcatag agccggatgt aaaagtcaga cctattaagg aagtggcccc 15060
cgtcttggg gtgcaaacgg tggacattca aatccccgtc accaccgctt caaccgccc 15120
ggaagctatg gaaacgaaa cggagacccc tgccgcgac ggtaccaggg aagtggcgct 15180
gcaaacggag ccttggtagc aatcgcagc ccctcggcgt cagaggcgtt cgcctcgta 15240
cggccccgcc aacgccatca tgccagaata tgcgctgcat ccgtctatc tgccactcc 15300

cggataccgg ggtgtgacgt atcgcccgtc tggaaccgc cgccgaaccc gtcgcgcgcg 15360
 ccgctcccgt cgcgctctgg cccccgtgtc ggtgcggcgt gtgaccgcgc ggggaaagac 15420
 agtcgtcatt cccaaccgc gttaccaccc tagcatcctt taataactct gccgttttgc 15480
 agatggctct gacttgccgc gtgcgccttc ccgttccgca ctatcgagga agatctcgtc 15540
 gtaggagagg catgacgggc agtggtcgcc ggcgggcttt gcgcaggcgc atgaaaggcg 15600
 gaattttacc cgccctgata cccatattg ccgcgcgccat cggtgccata cccggcgctg 15660
 cttcagtggc gttgcaagca gtcgtaata aataaacaag ggcttttgca cttatgacct 15720
 ggtcctgact attttatgca gaaagagcat ggaagacatc aattttacgt cgctggctcc 15780
 gcggcacggc tcggggccgc tcatgggcac ctggaacgac atcggcacca gtcagctcaa 15840
 cggggggcgt ttcaattggg ggagcctttg gagcggcatt aaaaactttg gctccacgat 15900
 taaatcctac ggcagcaaag cctggaacag tagtgctggc cagatgctcc gagataaact 15960
 gaaggacacc aacttccaag aaaaagtggc caatgggggt gtgaccggca tccacggcgc 16020
 ggtagatctc gccaaccaag cgggtgcagaa agagattgac aggcgtttgg aaagctcgcg 16080
 ggtgcgcgcg cagagagggg atgaggtgga ggtcgaggaa gtagaagtag aggaaaagct 16140
 gccccgctg gagaaagttc ccggtgcgcc tccgagaccg cagaagcggc ccaggccaga 16200
 actagaagag actctggtga cggagagcaa ggagcctccc tcgtacgagc aagccttgaa 16260
 agagggcgcc tctccacct cctaccgat gactaagccg atcgaccca tggctcgacc 16320
 ggtgtacggc aaggattaca agcccgctac gctagagctg cccccaccgc cccccacgcg 16380
 ccgacgcgc cccccctgc cgaactccgc ggcggccgcg gcgggaccgc tgtccgcacc 16440
 atccgctgtg cctctgcag cgcgccgtcc agtggccgtg gccactgcca gaaacccgag 16500
 aggccagaga ggagccaact ggcaaacgac gctgaacagc atcgtgggccc tgggagtgaa 16560
 aagcctgaaa cgcgcgcgtt gctattatta aaaaagtgtg gctaaaaagt ctcccgttgt 16620
 atacgcctcc tatgttaccg ccagagacga gtgactgtcg ccgcgagcgc cgctttcaag 16680
 atggccaccc catcgatgat gccgcagtgg tcttacatgc acatcgccgg ccaggacgcc 16740
 tcggagtacc tgagtcccg cctcgtgcag tttgcccgc ccaccgacac ctacttcagc 16800
 ttgggaaaca agtttagaaa cccacccgtg gccccaccc acgatgtgac cacggaccgc 16860
 tcgcagaggc tgaccctgcg ctttgtgccc gtagaccggg aggacaccgc gtactcttac 16920
 aaagtgcgct acacgttggc cgtaggggac aaccgagtgc tggacatggc cagcacctac 16980
 tttgacatcc ggggggtgct ggatcggggc cccagcttca agccctatc cggcaccgct 17040

tacaactccc tggcccccaa gggagctccc aaccctcgg aatggacgga cacttccgac 17100
aacaactta aagcatatgc tcaggctccc taccagagtc aaggacttac aaaggatggt 17160
attcagggtg ggctagttgt gacagagtca ggacaaacac cccaatatgc aaacaaagtg 17220
taccaacccg agccacaaat tggggaaaac caatggaatt tagaacaaga agataaagcg 17280
gcggaagag tcctaaagaa agataccct atgtttccct gctatgggtc atatgccagg 17340
cccacaaacg aacaaggagg gcaggcaaaa aaccaagaag tagatttaca gttttttgcc 17400
actccgggcg acaccagaa cacggctaaa gtggtacttt atgctgaaa tgtcaacctg 17460
gaaactccag atactcactt agtgtttaaa ccgatgacg acagcaccag ttcaaaactt 17520
cttcttgggc agcaggctgc acctaacaga cccaactaca taggttttag agataatttt 17580
attggtttaa tgtactaaa tagcactgga aacatggcg tgctggccg acaggcttct 17640
caattgaatg ccgtagtoga cttgcaggac agaaacaccg agttgtccta ccaactgatg 17700
ctggacgcac tgggggatcg cagccgatat ttttcaatgt ggaatcaggc agtagacagc 17760
tatgaccag acgttagaat tatagaaaac cacggagtgg aagacgaact gccaaactat 17820
tgttttctc tgggaggaat ggtggtgact gacaattaca actctgtgac gcctcaaaat 17880
ggaggcagtg gaaatacatg gcaggcagac aatactacat ttagtcaaag aggagcgcag 17940
attggctccg gaaacatgtt tgccctggaa attaactac aggccaacct ctggcgcggc 18000
ttcttgtatt ccaatattgg gtgtatctt ccagactctc tgaaaatcac ccccgacaac 18060
atcacgtgc cagaaaacaa aadacttat cagtacatga acggtcgcgt aacgccacc 18120
gggctcatag acacctatgt aaacgtgggc gcgcgctggt ccccgatgt catggacagc 18180
attaaccct tcaaccacca ccgtaacgcg ggcttgcgct accgctccat gctcttgggc 18240
aacggccgtt atgtgccttt tcacattcag gtgccccaaa aattctttgc cattaaaaac 18300
ctgtgcttc tccccggttc ctatacctat gagtggaaact tccgcaagga tgtcaacatg 18360
atcctgcaga gctcgtggg taatgacctg cgagtggacg gggccagcat acgctttgac 18420
agcattaacc tgtatgcaa cttttttccc atggcccaca acacggcctc taccctggaa 18480
gccatgctgc gcaacgacac caatgaccag tccttcaacg actacctgtg cgcggtaac 18540
atgctgtacc ccacccccgc caacgccacc agcgtgccca tttctattcc ttctcggaac 18600
tgggctgcct tcaggggctg gagttttact cgcctcaaaa ccaaggagac tccctcgctg 18660
ggctccggtt ttgacccta ctttgtttac tccggtcca ttcctacct agatggcacc 18720

ttttacctca accacacttt caaaaagggtg tctattatgt ttgactcctc ggtagctgg	18780
cccggaacg accgcctgct aacgccaac gagtcgaaa ttaagcggtc cgtggacggt	18840
gaagggtaca acgtggccca gagcaacatg accaaggact ggtttctaata tcaaagtctc	18900
agtcactata atatagggtta ccagggttc tatgtgcccg agaactacaa ggaccgcatg	18960
tactccttct tccgcaactt ccaaccaatg agccggcagg tggtagatac cgtgacttat	19020
acagactaca aagatgtcaa gctccctac caacacaaca actcagggtt cgtgggctac	19080
atgggaccca ccatgcgaga gggacaggcc taccggcca actatcccta cccctgatc	19140
ggagagactg ccgtaccag cctcacgcag aaaaagttcc tctgcgaccg ggtgatgtgg	19200
aggataccct tctctagcaa ctttatgtcg atgggtccc tcaccgacct ggggcagaac	19260
atgctgtacg ccaactccgc tcaggccttg gacatgactt ttgagggtga tcccatggat	19320
gagcccacgc ttctctatgt tctgtttgaa gtcttcgacg tggtagcat ccaaccagccg	19380
caccgcggcg tcatcgagge cgtctacctg cgcacacctt tctctgccgg taacgccacc	19440
acctaagaa gctgatgggt tccagcgaac aggagtgtga ggccattgtt cgcgacctgg	19500
gctgcgggccc ctgctttttg ggcaccttcg acaagcgttt tcccggttc atgtcccccc	19560
acaagccggc ctgcgccatc gttaacacgg ccggacggga gacagggggg gtgcactggc	19620
tgccttcgc ctggaacccg cgcaaccgca cctgctacct gttagacctt ttgggtttct	19680
ccgacgaaag gctgaagcag atctaccaat tcgagtacga ggggtcctc aagcgcagcg	19740
ctctggcctc cagccccgac cactgcgtca ccctggaaaa gtccaccag acggtccagg	19800
ggccccctc ggccgctgc gggcttttct gttgcatgtt ttgacagcc ttctgtcact	19860
ggcctcacac ccccatggag cgcaaccca ccatggatct gtcaccgga gtgccaaca	19920
gcatgcttca cagtccccag gtgcgccccca ccctgcgtcg caatcaggac cacctgtatc	19980
gctttctggg gaaacactct gcctatttcc gccgccaccg gcagcgcac gaacaggcca	20040
cggccttcga aagcatgagc caaagagtgt aatcaataaa aaccgttttt atttgacatg	20100
atacgcgctt ctggcgtttt tattaaaaat cgaagggttc gagggagggg tctctgtgcc	20160
cgctggggag ggacacgttg cggtagtga atcgggcgct ccaacgaaac tcggggatca	20220
ccagccgagg cagggccacg tcttccatgt tctgcttcca aaactgtcgc accagctgca	20280
gggtcccat cagtcgggc gctgagatct tgaagtcgca gttagggccg gagccccgc	20340
ggctgttgcg gaacacgggg ttggcacact ggaacaccaa cagctgggg ttgtggatac	20400
tagccagggc cgtcgggtcg gtcacctccg atgcatccag atcctcgga ttgctcaggg	20460

cgaacggggt cagcttgac atctgccgcc cgatctgggg taccaggctg cgcttggtga 20520
ggcagtcgca gcgcagaggg atgaggatgc gacgctgccc gcgttgcatg atggggtaac 20580
tcgccgccag gaactcctct atctgacgga aggccatctg ggccttgacg ccttcggtga 20640
aaaatagccc acaggacttg ctggaaaaca cgttattgcc acagttgatg tcttcgcgcg 20700
agcagcgcg atcttcggtt ttcagctgaa ccacgttgcg accccagcgg ttctgaacca 20760
ccttggtttt cgtgggatgc tccttcagcg cccgctgtcc gttctcgctg gtcacatcca 20820
tttcaccac gtgctccttg cagaccatct ccactccgtg gaaacagaac agaatgccct 20880
cctgttgggt attgcgatgc tcccacacgg cgcacccggg ggactcccag ctcttggtgt 20940
tcacccccgc gtaggcttcc atgtaagcca ttagaaatct gcccatcagc tcagtgaagg 21000
tcttctggtt ggtgaagggt agcggcaggg cgcgggtgtt ctcgttcaac caagtttgac 21060
agatcttgcg gtacacgggt ccttggtcgg gcagaaactt aaaagtcgtt ctgctctcgt 21120
tgtccacgtg gaacttctcc atcaacatcg tcatgacttc catgcccttc tcccaggcag 21180
tcaccagcgg cgcgctctcg gggttcttca ccaacacggc ggtggagggg ccttcgcccg 21240
ccccgacgtc cttcatggac attttttgaa actccacggg gccgtccgcg cggcgctactc 21300
tgcgcatcgg agggtagctg aagcccacct ccatgacggg gctttcgccc tcgetgtcgg 21360
agacgatctc cggggagggc ggcggaacgg ggcgagactt gcgagccttc ttcttgggag 21420
ggagcggagg cacctcctgc tcgcgctcgg gactcatctc ccgcaagtag ggggtgatgg 21480
agcttcctgg ttggttctga cggttggcca ttgtatccta ggcagaaaga catggagctt 21540
atgcgcgagg aaactttaac cgccccgtcc cccgtcagcg acgaagaggt catcgtcgaa 21600
caggaccggg gctacgttac gccgcccag gatctggagg ggcccttaga cgaccggcgc 21660
gacgctagtg agcggcagga aaatgagaaa gaggaggagg agggctgcta cctcctggaa 21720
ggcgacgttt tgctaaagca ttctgccagg cagagcacca tactcaagga ggccttgcaa 21780
gaccgctccg aggtgccctt ggacgtcgcc gcgctctccc aggcctacga ggcgaacctt 21840
ttctcgcccc gagtgccctc gaagagacag cccaacggca cctgcgagcc caaccgcga 21900
ctcaacttct acccgtgtt cgcggtgccc gaggcgctgg ccacctacca catcttttct 21960
aaaaaccagc gcattccctt ttcttgccgg gccaaaccga ccgcgccga taggaagcta 22020
acactcagaa acggagtcag catacctgat atcacgtcac tggaggaagt gcctaagatc 22080
ttcgagggtc tgggtcgaga tgagaagcgg gcggcgaacg ctctgcagaa agaacagaaa 22140

gagagtcaga acgtgctggt ggagctggag ggggacaacg cgcgtctgac cgtcctcaaa 22200
cgttgcatag aagtttccca ctctgcctac ccggccctca acctgccgcc caaagttagt 22260
aaatcgggtca tggaccagct actcatcaag agagctgagc cctgaatcc cgaccacctt 22320
gagggcgaaa actcagagga cggaaagccc gtcgtcagcg acgaggagct cgagcgggtg 22380
ctggaaacca gggaccccca gcagtggcaa gagaggcgca agatgatgat ggcggccgtg 22440
ctggtcacgg tggagctaga atgcctgcaa cggtttttca gcgacgtgga gacgctacgc 22500
aaaatcgggg agtccctgca ctacaccttc cgccagggtt acgttcgccca ggctgcaaa 22560
atctccaacg tagagctcag caacctgggtt tcctacatgg gcctcctcca cgagaaccgg 22620
ctggggcaga gcgtgctgca ctgcaccttg caaggcgagg cgcgaaggga ctacgtccga 22680
gactgcgtct acctcttctt cacctcacc tggcagaccg ccatgggcgt gtggcagcag 22740
tgcttgggaag agagaaacct caaagagctg gacaaactcc tctgccgccca gcggcgggcc 22800
ctctggaccg gcttcagcga gcgcacggtc gcctgcgccc tggcagacat cattttccca 22860
gaacgcctga tgaaaacctt gcagaacggc ctgccggatt tcatcagtcga gagcatcttg 22920
caaaacttcc gctcctctgt cctggagcgc tccgggatct tgcccgccat gagctgcgcg 22980
ctgccttctg actttgtccc cctttcctac cgcgagtgcc cccccact gtggagccac 23040
tgctacctct tccaactggc caactttctg gcctaccact ccgacctcat ggaagacgtg 23100
agcggagagg ggctgctcga gtgccactgc cgctgcaacc tctgcacccc ccacagatcg 23160
ctggcctgca acaccgagct gctcagcgaa acccagggtca taggtacctt cgagatccag 23220
gggcccagc agcaagagggt tgcttcgggc ttgaagctca ctccggcgct gtggacctcg 23280
gcttacttac gcaaatttgt agccgaggac taccacgccc acaaaattca gttttacgaa 23340
gaccaatctc gaccaccgaa agctcccttc acggcctgcg tcatcaccca gagcaaaatc 23400
ctggcccaat tgcaatccat caaccaagcg cgccgagatt tccttttgaa aaagggtcgg 23460
ggggtgtacc tggaccccca gaccggcgag gaactcaacc cgtccacact ttccgtcgaa 23520
gcagccccc cgagacatgc caccgaagggt aaccgccaaag cagctgatcg ctccggcagag 23580
agcgaagaag caagagctgc tccagcagca ggtggaggac gaggaagagc tgtgggacag 23640
ccaggcagag gaggtgtcag aggacgagga ggagatggaa agctgggaca gcctagacga 23700
ggaggacgag ctttcagagg aagaggcgac cgaagaaaaa ccacctgcat ccagcgcgcc 23760
ttctctgagc cgacagccga agccccggcc cccgacgccc ccggccggct cactcaaagc 23820
cagccgtagg tgggacgcca ccggatctcc agcggcagcg gcaacggcag cgggttaaggc 23880

caaacgcgag cggcgggggt attgctcctg gcggaccac aaaagcagta tcgtgaactg 23940
 cttgcaacac tgccggggaa acatctcctt tgcccgacgc tacctcctct tccatcacgg 24000
 tgtggccttc cctcgcaacg ttctctatta ttaccgtcat ctctacagcc cctacgaaac 24060
 gctcggagaa aaaagctaag gcctcctctg ccgcgaggaa aaactccgcc gccgctgccg 24120
 ccaaggatcc gccggccacc gaggagctga gaaagcgcct ctttcccact ctgtatgcta 24180
 tctttcagca aagccgcggg cagcaccctc agcgcgaact gaaaataaaa aaccgctcct 24240
 tccgctcact caccgcgacg tgtctgtacc acaagagaga agaccagctg cagcgacccc 24300
 tggacgacgc cgaagcactg ttcagcaaact actgctcagc gtctcttaaa gactaaaaga 24360
 cccgcgcttt tttcccctcg ggcgcacaaa cccacgtcat cgccagcatg agcaaggaga 24420
 ttcccacccc ttacatgtgg agctatcagc cccagatggg cctggccgcg ggggcccgcc 24480
 aggactactc cagcaaaatg aactggctca gcgcggcccc ccacatgatc tcacgagtta 24540
 acggcatccg agcccaccga aaccagatcc tcttagaaca ggcggcaatc accgccacac 24600
 cccggcgcca actcaaccgg ccaggttggc ccgcgcacca ggtgtatcag gaaactcccc 24660
 gcccgaccac agtcctcctg ccacgcgacg cggaggccga agtcctcatg actaactctg 24720
 ggggtacaatt agcggggcggg tccaggtacg ccaggtacag aggtcgggccc gctccttact 24780
 ctcccgggag tataaagagg gtgatcatc gaggcggagg tatccagctc aacgacgagg 24840
 cgggtgagctc ctcaaccggg ctccagacctg acggagtctt ccagctcggg gggcggggcc 24900
 gctcttcctt caccactcgc caggcctacc tgacctgca gagctcttcc tcgcagccgc 24960
 gctccggggg aatcggcact ctccagttcg tggaagagtt cgtcccctcc gtctacttca 25020
 acccgtttc cggctcaact ggacgctacc cggacgcctt cattcccaac tttgacgcag 25080
 tgagtgaatc cgtggacggc tacgactgat gacagatggg gcggccgtga gagctcggct 25140
 gcgacatctg catcactgcc gccagcctcg ctgctacgct cgggaggcga tcgtgttcag 25200
 ctactttgag ctgccggacg agcaccctca gggaccggct cacgggttga aactcgagat 25260
 tgagaacgcg cttgagtctc acctcatcga cgccttcacc gcccggcctc tcctggtaga 25320
 aaccgaacgc gggatcacta ccatcaccct gttctgcate tgccccacgc ccggattaca 25380
 tgaagatctg tggtgtcatc tttgcgctca gtttaataaa aactgaactt tttgccgtac 25440
 cttcaacgcc acgcgttggt tctccttggt aaaaaacccc aggagtcctt aacttacaca 25500
 tagcaaaacc cttgtatttt accatagaaa aacaactagc cctttcaatt ggaaaagggt 25560

taacaatttc tgctacagga cagttggaaa gcacagcaag cgtacaggac agcgctacac	25620
caccctacg tggatattcc cctttaaagc tgacagacaa cggtttaaca ttaagctatt	25680
cagatcccct gcgtgtggta ggtgaaccaac ttacgtttaa ttttacttct ccactacgtt	25740
acgaaaatgg cagtcttaca ttcaactaca cttctcccat gacactaata aacaacagtc	25800
ttgctattaa cgtcaatacc tccaaggcc tcagtagtga caacggcaca ctcgctgtaa	25860
atgttactcc agattttaga tttaacagct ctggcgcctt aacttttggc atacaaagtc	25920
tatggacttt tccaaccaa actcctaact gtaccgtgtt taccgaaagt gactccctgc	25980
tgagtctttg cttgactaaa tgcggagctc acgtacttgg aagcgtgagt ttaagcggag	26040
tggcaggaac catgctaaaa atgaccacaca cttctgttac cgttcagttt tcgtttgatg	26100
acagtggtaa actaatattc tctccacttg cgaacaacac ttgggggtgtt cgacaaagcg	26160
agagtccgtt gcccaacca tcttcaacg ctctcagctt tatgccaaac agtacattt	26220
attctagagg agcaagtaac gaacctcaaa acaattatta tgtccagacg tatcttagag	26280
gcaacgtgcg aaagccaatt ctactaactg ttacctacaa ctcagttaat tcaggatatt	26340
ccttaacttt taaatgggat gctgtcgcca atgaaaaatt tgccactcct acatcttcgt	26400
tttgctatgt tgcagagcaa taaaaccctg ttaccccacc gtctcgtttt tttcagatga	26460
aacgagcgag agttgatgaa gacttcaacc cagtgtaccc ttatgacccc ccatacgctc	26520
ccgtcatgcc cttcattact ccgcctttta cctcctcgga tgggttgacg gaaaaaccac	26580
ttggagtgtt aagtttaaac tacagggatc ccattactac acaaaaatggg tctctcacgt	26640
taaaactagg aaacggcctc actctaaaca accagggaca gttaacatca actgctggcg	26700
aagtggagcc tccgctcact aatgctaaca acaacttgc actagcctat agcgaaccat	26760
tagcagtaaa aagcaaccgc ctaactctat cacacacgcg tccccttgtc atcgctaata	26820
attcttttagc gttgcaagtt tcagagccta tttttgtaaa tgacgatgac aagctagccc	26880
tgacagacgc cgccccctt gtaaccaacg ctggcaccct tcgcttacag agcgctgccc	26940
ctttaggatt ggttgaaaat actcttaaac tgctgttttc taaacccttg tatttgcaaa	27000
atgattttct tgcattagcc attgaacgcc ccctggctgt agcagccgca ggtactctga	27060
ccctacaact tactcctcca ttaaagacta acgatgacgg gctaactacta tccaagtcg	27120
agccattaac tgtaaaaaac ggaaacctag gcttgcaaat atcgcgccct ttagttgttc	27180
aaaacaacgg cctttcgctt gctattaccc ccccgctgcg tttgtttaac agcgaccccg	27240
ttcttggttt gggcttcaact tttcccttag ctgtcacaaa caacctctc tccttaaaca	27300

tgggagacgg agttaaactt acctataata aactaacagc caatttgggt agggatttac 27360
 aatttgaaaa cgggtgcgatt gccgtaacgc ttactgccga attacctttg caatacacta 27420
 acaaacttca actgaatatt ggagctggcc ttcgttaca tggagccagc agaaaactag 27480
 atgtaaacat taaccaaaat aaaggcttaa cttgggacaa cgatgcagtt attcccaaac 27540
 taggatcggg cttacaattt gaccctaatt gcaacatcgc tgttatccct gaaaccgtga 27600
 agccgcaaac gttatggacg actgcagatc cctcgcctaa ctgctcagt taccaggact 27660
 tggatgccag gctgtggctc gctcttgta aaagtggcga catgggtgcat ggaagcattg 27720
 ccctaaaagc cctaaaaggg acgttgctaa atcctacagc cagctacatt tccattgtga 27780
 tatattttta cagcaacgga gtcaggcgta ccaactatcc aacgtttgac aacgaaggca 27840
 ccttagctaa cagcgccact tggggatacc gacaggggca atctgctaac actaatgtga 27900
 ccaatgccac tgaatttatg ccagctcaa gcaggtagcc cgtgaataaa ggagacaaca 27960
 ttcaaaatca atctttttca tacacctgta ttaaaggaga ttttgctatg cctgtcccgt 28020
 tccgtgtaac atataatcac gccctggaag ggtattccct taagtccacc tggcgcggtg 28080
 tagccaatca ggcccttgat attccttgct gttcattttc atacatcaca gaataaaaaa 28140
 ccactttttc attttaattt ctttttattt tacacgaaca gtgagacttc ctccaccctt 28200
 ccatttgaca gcatacacca gcctctcccc cttcatagca gttaaactgtt gtgaatcagt 28260
 cccgtatttg ggagttaaaa tccaacagt ctctttgggtg atgaaacgtc gatcagtaat 28320
 ggacacaaat ccctgggaca ggttttccaa cgtttcgggtg aaaaactgca caccgccccta 28380
 caaaacaaac aggttcaggc tctccacggg ttatctcccc gatcaaaactc agacagggtg 28440
 aagggtgcggg ggtgttcac taaaccacgc aggtggcgct gtctgaacct ctcggtgcca 28500
 ctctgtgag gctggtaaga agttagattg tcdagtagcc tcacagcatg tatcatcagt 28560
 ctacgagtgc gtctggcgca gcagcgcatc tgaatctcac tgagattccg gcaagaatcg 28620
 cacaccatca caatcagggtt gttcatgatc ccatagctga acacgctcca gccaaagctc 28680
 attcgctcca acagcgccac cgcgtgtccg tccaacctta ctttaacata aatcagggtg 28740
 ctgccgcgta caaacatgct acccacatac agaacttccc ggggcaggcc cctgttcacc 28800
 acctgtctgt accagggaaa cctcacattt atcaggggagc catagatggc cattttaaac 28860
 caattagcta ataccgcccc accagctcta cactgaagag aaccgggaga gttacaatga 28920
 cagtgaataa tccatctctc ataaccctg atggtctgat gaaaatctag atctaacgtg 28980

gcacaacaaa tacacacttt catatacatt ttcataacat gtttttccca ggccgttaaa 29040
atacaatccc aatacacggg ccactcctgc agtacaataa agctaataca agatggtata 29100
ctcctcacct cactgacact gtgcatgttc atattttcac attctaagta ccgagagttc 29160
tcctctacag cagcactgct gcggtcctca caaggtggta gctggtgatg attgtagggg 29220
gccagtctgc agcgataccg tctgtcgcgt tgcacgtag accaggaacc gacgcacctc 29280
ctcgtacttg tggtagcaga accacgtccg ctgccagcac gtctccacgt aacgcgggtc 29340
cctgcgtcgc tcacgtcccc tcctcaatgc aaagtgaac cactcttgta atccacacag 29400
atccctctcg gcctccgggg tgatgcacac ctcaaacctc cagatgtctc ggtacagtcc 29460
caaacacgta gtgagggcga gttccaacca agacagacag cctgatctat ccgacacac 29520
tgaggtgga ggaagacacg gaagaggcat gttattccaa gcgattcacc aacgggtcga 29580
aatgaagatc ccgaagatga caacgggtcgc ctccggagcc ctgatggaat ttaacagcca 29640
gatcaaacgt tatgcgattc tccaagctat cgatcgccgc ttccaaaaga gcctggaccc 29700
gcacttcac aaacaccagc aaagcaaaag cactattatc aaactcttca atcatcaagc 29760
tgcaggactg tacaatgcct aagtaatttt cgtttctcca ctgcggaatg atgtcgcggc 29820
agatagtctg aagggtcatc ccgtgcaggg taaaagctc cgaaagggcg ccctctacag 29880
ccatgcgtag acacaccatc atgactgcaa gatatcgggc tcctgagaca cctgcagcag 29940
atttaacaga tcaaggctcag gttgctctcc gcgatcacga atctccatcc gcaaggctcat 30000
ttgcaaaaaa tttaataaat ctatgccgac tagatctgtc aactccgcat taggaaccaa 30060
atcagggtgtg gctacgcagc acaaaagttc cagggatggt gccaaactca ctagaaccgc 30120
tcccgagtaa caaaactgat gaatgggagt aacacagtgt aaaatgtgca accaaaaatc 30180
actaagggtgc tccttttaaaa agtccagtac ttctatatte agtccgtgca agtactgaag 30240
caactgtgcg ggaatatgca caacaaaaaa aatagggcgg ctgagataca tggtgacctc 30300
aaataaaaag aatcattaaa ctaaaagaagc ttggcgaacg gtgggataaa tgacacgctc 30360
cagcagcaga caggcaaccg gctgtcccgc ggaaccgcgg taaaatcat ccgaatgatt 30420
aaaaagaaca acagaaactt cccaccatgt actcggttgg atctcctgag cacacagcaa 30480
tacccccctc acattcatgt ccgccacaga aaaaaaacgt ccagatacc cagcgggggat 30540
atccaacgac agctgcaaag acagcaaac aatccctctg ggagcgatca caaaatcctc 30600
cggtgaaaaa agcacataca tattagaata accctgttgc tggggcaaaa aggccccggc 30660
tcccagcaaa tgcacataaa tatgttcac agccattgcc ccgtcttacc gcgtaatcag 30720

ccacgaaaaa atcgagctaa aattcaccac acagcctata gctatatata cactccgccc 30780
aatgacgcta ataccgcacc acccaccgacc aaagttcacc cacaccacaca aaacccgcga 30840
aaatccagcg ccgtcagcac ttccgcaatt tcagtctcac aacgtcactt ccgcgcgcct 30900
tttcacattc ccacacacac ccgcgcctt cgcgccgccc tcgcgccacc ccgcgtcacc 30960
gcacgtcacc ccggccccgc ctgcgtcctc cccgtcatt atcatattgg cacgtttcca 31020
gaataaggtg tattattgat gatg 31044

<210> 6
<211> 34115
<212> DNA
<213> simian adenovirus SV-39

<400> 6
catcatcaat ataacaccgc aagatggcga ccgagttaac atgcaaatga ggtgggcgga 60
gttacgcgac ctttgtcttg ggaacgcgga agtgggcgcg gcgggtttcg gggaggagcg 120
cggggcgggg cgggcgtgtc gcgcggcggt gacgcgccgg ggaccggaa attgagtagt 180
ttttattcat ttgtcaagtt tttctgtaca ttttggcgcg aaaactgaaa cgaggaagtg 240
aaaagtga aaatgccgagg tagtcaccgg gtggagatct gacctttgcc gtgtggagtt 300
taccgctga cgtgtgggtt tcggtctcta ttttttact gtgttttcc gggtagcgtc 360
aaaggtcccc attttatgac tccacgtcag ctgatcgcta gggattttaa tgcgcctcag 420
accgtcaaga ggccactctt gagtgcgggc gagaagagtt ttctctccg cgttccgcca 480
actgtgaaaa aatgaggaac ttcttgctat ctccggggct gccagcgacc gtagccgccg 540
agctgttga ggacattgtt accggagctc tgggagacga tcctcaggtg atttctcact 600
tttgtgaaga ttttagtctt catgatctct atgatattga tccgggtgtt gaggggcaag 660
aggatgaatg gctggagtct gtggatgggt tttttccgga cgctatgctg ctagaggctg 720
atttgcacc acctcacaac tctcacactg agcccgagtc agctgctatt cctgaattgt 780
catcagggtga acttgacttg gcttggttac agactatgcc tccggagtcg gatgaggagg 840
acagcgggat cagcgatccc acggctttta tggctctctaa ggcgattgct atactaaaag 900
aagatgatga tggcgatgat ggatttcgac tggacgctcc ggcggtgccg gggagagact 960
gtaagtccctg tgaataccac cgggatcgta ccggagaccc gtctatgttg tgttctctgt 1020
gttatctccg tcttaacgct gcttttctct acagtaagtg ttttctgctt ttttaccctg 1080
tggctttgtt gagtttattt ttttctgtgt ctcatagggt gttgtttatt ataggctcctg 1140

tttcagatgt ggaggaacct gatagtacta ctggaaatga ggaggaaaag ccctccccgc	1200
cgaaactaac tcagcgtgc agacctata ttttgagacc ctgggccag cgtgtgtcat	1260
cccgaaacg tgctgctgtt aattgcatag aagatttatt ggaagagccc actgaacctt	1320
tggacttgtc cttaaagcga ccccgccgc agtagggcgc ggtgccagtt tttctctct	1380
agcttccggg tgactcagt caataaaaat tttcttgga acaggtgtat gtgtttactt	1440
tacgggcggg aagggattag gggagtataa agctggagg gaaaaatctg aggctgtcag	1500
atcgagttag aagttccatg gacttgtacg agagcctaga gaatctaagt tctttgcgac	1560
gtttgctgga ggaggcctcc gacagaacct cttacatttg gaggtttctg ttcggttccc	1620
ctctgagtcg ctttttgac cggtgaagc gagagcacct gacggaattt gatgggcttt	1680
tagagcagct gcctggactg tttgattctt tgaatctcg ccaccggacg ctgctagagg	1740
agaggctttt tccacaattg gacttttct ctccaggccg tctgtgttca gcgcttgctt	1800
ttgctgtaca tctgttgga agatggaacg agcagacgca gctcagccc gggtacactc	1860
tggacttcct gacgctatgc ctatggaagt tcggaatcag gagggggagg aagctgtacg	1920
ggcgcttggg ggagaggcat ccgtctctgc gccagcagcg tctgcaagct caagtgtgc	1980
tgaggcgga ggatctgga gccatttcgg agggaggag cgcatgga gagaagaatc	2040
cgagagcggg gctggacct ccggcgagg agtaggggg ataccggacc cttttcctga	2100
gttggtcttg gggcggttg gggcgcttc tgtggtacgt gaggatgaag agggcgcca	2160
acgggtcag aagagggagc attttgagtc ctgactttc ttggctgatg taaccgtggc	2220
cctgatggcg aaaaacaggc tggaggtgg gtggtacct gaagtatgg aggactttga	2280
gaagggggac ttgcacctgc tggaaaaata taactttgag caggtgaaa catactggat	2340
gaaccggat gaggactgg aggtggttt gaaccgatac ggcaaggtag ctctgcgtcc	2400
cgactgtgc taccaggtc gcgacaagg ggtcctgcga cgcaacgtgt acctgttggg	2460
caacggcgcc accgtggaga tggtggaacc cagaaggggt ggttttgtgg ccaatatgca	2520
agaaatgtgc cctggggtg tgggttgct tgggtgact tttcatagt tgaggtttag	2580
cggtagcaat tttgggggtg tggttattac cggaacact cctgtggtcc tgcataattg	2640
ctactttttt ggcttcagca acacctgtgt ggaaatgagg gtgggaggca aagtgcgcgg	2700
gtgttccttt tacgcttgc ggaaggggt ggtgagccag ggtaaggcta aagtgtctgt	2760
tcacaagtgt atgttgaga gatgcacct gggcatttc agtgagggt tcctccacgc	2820

cagcgacaac gtggcttctg acaacggctg cgcctttctt atcaagggag ggggtcgcat	2880
ctgtcacaac atgatatgcg gccctgggga tgtccccca aagccttacc agatgggttac	2940
ctgcacagat ggcaagggtgc gcatgtctaa gcctgtgcac attgtgggcc accggcgcca	3000
ccgctggcca gagtttgaac acaatgtgat gaccgctgt agcttgtacc tgggaggcag	3060
gcgaggagtt ttcttgccca gacagtgtaa cctggccac tgcaacgtga tcatggaaca	3120
atccgccgct acccaggttt gctttggagg aatatttgat ataagcatgg tgggtgataa	3180
gacctgcgc tacgacgact gtccgggctcg tactcgaacc tgcgactgcg gacccctca	3240
cctgtgtaac ctgactgtga tggggatggt gactgaggag gtgcgactgg accactgtca	3300
gcactcttgc ctgccccagg agttttcttc ctccgacgag gaggactagg taggtggttg	3360
gggcgtggcc agcgagaggg tgggctataa aggggagggtg tcggctgacg ctgtcttctg	3420
tttttcaggt accatgagcg gatcaagcag ccagaccgcg ctgagcttcg acggggccgt	3480
gtacagcccc tttctgacgg ggcgcttgcc tgcctgggcc ggagtgcgtc agaattgttac	3540
cggttcgacc gtggacggac gtcccgtgga tccatctaac gctgcttcta tgcgctacgc	3600
tactatcagc acatctactc tggacagcgc cgctgcgcgc gcagccgcca cctcagccgc	3660
tctctccgcc gccaatgca tggctattaa cccaagcctt tacagccctg tatccgtgga	3720
cacctcagcc ctggagcttt accggcgaga tctagctcaa gtggtggacc aactcgcagc	3780
cgtgagccaa cagttgcagc tgggtgcgac ccgagtggag caactttccc gccctcccca	3840
gtaaccgcaa aaattcaata aacagaatct aataaacagc acttgagaaa agtttaact	3900
tgtggttgac ttatttctg gatagctggg gggagggaac ggcgggaacg gtaagacctg	3960
gtccatcggt cccggtcggt gagaacacgg tggatttttt ccaagaccgc atagagggtg	4020
gtctgaacgt tgagatacat gggcatgagc cctgtctcggg ggtggaggta ggccactgc	4080
agggcctcgt tttcaggggt ggtgttgtaa atgatccagt cgtaggcccc ccgctgggcg	4140
tgggtgctgga agatgtcctt cagcagcaag ctgatggcaa cgggaagacc cttggtgtag	4200
gtgttgacaa agcgggtgag ttgggagggg tgcattgcgg gactgatgag gtgcattttg	4260
gcctggatct tgagggtggc tatgttgccg ccagatcgc gcctgggatt catgttatgc	4320
aagaccacca gcaccgagta accggtgcag cgggggaatt tgcgtgcag cttggaaggg	4380
aaagcgtgga agaatttgga gaccctcgg tgcgcgcta ggttttccat gcactcatcc	4440
atgatgatgg cgatgggccc ccgggaggca gcctgggcaa aaacgttgcg ggggtccgtg	4500
acatcgtagt tgtggtcctg ggtgagttca tcataggaca ttttgacaaa gcgcgggcag	4560

aggggtccag actggggaat gatgggtcca tccgggtccg gggcgtagtt gccctcgag 4620
 atttgcattt cccaggcttt gatttcagag ggagggatca tgtcaacctg gggggcgatg 4680
 aaaaaaatgg tctctggggc gggggtgatg agctgggtgg aaagcagggt gcgcaagagc 4740
 tgtgacttgc cgcagccggt gggcccgtag atgacagcta tgacgggttg cagggtgtag 4800
 tttagagagc tacaactgcc atcatccttc aaaagcgggg ccacactggt taaaagttct 4860
 ctaacatgta agttttcccg cactaagtcc tgcaggagac gtgacctcc tagggagaga 4920
 agttcaggaa gcgaagcaaa gtttttaagt ggcttgaggc catcgccaa gggcaagttc 4980
 ctgagagttt gactgagcag ttccagccg tccagagct cggttacgtg ctctacggca 5040
 tctcgatcca gcagacctcc tcgtttcggg ggttggggcg gctctggctg tagggaatga 5100
 ggcggtgggc gtccagctgg gccatgggtc ggtccctcca tgggcgcagg gttctcttca 5160
 ggggtggtct ggtcacgggt aatgggtggg ccccgggctg ggcgctggcc aggggtgcgt 5220
 tgaggctgag gcggtgggt gcgaaccgtt gcttttcgtc tccctgcaag tcagcctaat 5280
 agcaacggac catgagctca tagtccaggc tctctcgggc atgtcctttg gcgcgaagct 5340
 tgcctttgga aacgtgccc cagtttgagc agagcaagca ttttagcgcg tagagttttg 5400
 gcgccaagaa cacggattcc ggggaataag catccccacc gcagttggag caaacggttt 5460
 cgcattccac cagccaggtc agctgaggat cttttgggtc aaaaaccaag cgcccgccgt 5520
 tttttttgat gcgcttccca cctcgggtct ccatgaggcg gtgcccgcgt tcggtgacga 5580
 agaggctgtc ggtgtctccg tagacggagg tcagggcgcg ctccctccagg ggggtcccgc 5640
 ggtcctcggc gtagagaaac tcgcaccact ctgacataaa cgcccggtc caggctagga 5700
 cgaatgaggc gatgtgggaa gggtaaccgt cgttatcgat gagggggtcg gttttttcca 5760
 aggtgtgcag gcacatgtcc ccctcgtccg cttccaaaaa tgtgattggc ttgtaggtgt 5820
 aagtcacgtg atcctgtcct tccgcggggg tataaaaggg ggcgtttccc cctcctcgt 5880
 cactctcttc cggttcgtg tcgccaagg ccagctgttg gggtagctaa acgcggtga 5940
 aggcgggcat gacctgtgcg ctgaggttgt cagttctat atacaggaa gatttgatgg 6000
 cgagcgcccc cgtggagatg cccttgaggt gctcggggcc catttggta gaaaacacaa 6060
 tctgtcgggt atcaagcttg gtggcaaaag acccgtagag ggcgttgag agcaacttg 6120
 cgatggagcg ctgggtttgg ttttttccc ggtcggcttt ttcctggcc gcgatgtga 6180
 gctggacgta ctccctggcc acgcacttcc agccggga aaacggcgctg cgtcgtccg 6240

gcaccagcct	cacgctccat	ccgcggttgt	gcaggggtgat	gacgtcgatg	ctggtggcca	6300
cctctccg	caggggctcg	ttgggtccagc	agagggcgacc	gcccttgca	gagcagaagg	6360
ggggcagggg	gtcaagcagg	cgctcgctccg	gggggtcggc	gtcgatggta	aagatggcgg	6420
gcagcaggtg	tttgtcaaag	taatcgatct	gatgcccggg	gcaacgcagg	gcggtttccc	6480
agtcccgcac	cgccaaggcg	cgctcgatg	gactgagggg	ggcgccccag	ggcatgggat	6540
gcgtcagggc	cgaggcgta	atgcgcgaga	tgcatagac	gtaaaggggc	tcctccagga	6600
cgccgaggtg	gggtgggtag	cagcgcccc	cgcggtgct	ggcccgtag	tagtcgtaga	6660
gctcgtcgca	gggggcccaga	aggtggcgcc	tgaggtgagc	gcgctggggc	ttttcatctc	6720
ggaagaggat	ctgcctgaag	atggcggtggg	agttggagga	gatggtgggc	cgctgaaaa	6780
tggtgaagcg	ggcgctcgggc	agacccacgg	cctcgccgat	aaagtgggcg	taggactctt	6840
gcagcttttc	caccagggag	gcgggtgacca	gcacgtccag	agcgagtag	tccagggttt	6900
cccgacgat	gtcataatgc	tcttccctttt	tttccctcca	gaggtctcgg	ttgaagagat	6960
actcttcg	gtctttccag	tactcttgga	gaggaaaccc	gttttcgtct	ccacggtaag	7020
agcccaacat	gtaaaactgg	ttgacggcct	gatagggaca	gcaccccttc	tcacagggca	7080
gcgagtaggc	cagggcgggc	ttgcgcaggg	aggtgtgagt	cagggcaaa	gtgtcgcgga	7140
ccataacttt	tacaaactgg	tacttaaagt	cccggtcgtc	gcacatgcct	cgctcccagt	7200
ctgagtagtc	tgtgcgcttt	ttgtgcttgg	ggtaggcag	ggagtaggtg	acgtcgtaa	7260
agaggatttt	gccacatctg	ggcataaagt	tgcgagagat	tctgaagggg	ccgggcacct	7320
ccgagcggtt	gttgatgact	tgggcagcca	ggagaatttc	gtcgaagccg	ttgatgtgt	7380
gccccacgac	gtagaactct	atgaaacggc	gagcgccggc	cagcaggggg	cacttttcaa	7440
gttgctggaa	agtaagttcc	cgcggtcga	cgccgtgttc	cgtgcggctc	cagtcctcca	7500
ccgggtttcg	ctccacaaaa	tcctgccaga	tgtggtcgac	tagcaagagc	tcagtcgggt	7560
cgcgaaattc	gcggaatttt	ctgccgatgg	cttgcttctg	ggggttcaag	caaaaaagg	7620
tgtctgcgtg	gtcgcgcag	gcgtcccagc	cgagctcgcg	agccagattc	agggccagca	7680
gcaccagagc	cggtccaccg	gtgattttca	tgacgaggag	aaagggcacc	agctgttttc	7740
cgaacgcgcc	catccagggtg	taggtctcca	cgtcgtaggt	gagaaacaga	cgttcgggtcc	7800
gcgggtgcga	tcccaggggg	aaaaacttga	tgggctgcca	ccattgggag	ctctgggctg	7860
ggatgtgatg	gaagtaaaag	tcccggcggc	gcgtggaaca	ttcgtgctgg	tttttgtaaa	7920
agcgggccgca	gtggtcgag	cgcgagacgg	agtgaaggct	gtgaatcagg	tgaatcttgc	7980

gtcgtgagg	gggccccaga	gccaaaaagc	ggagcgggaa	cgaccgcgcg	gccacttcgg	8040
cgtccgcagg	caagatggat	gagggttcca	ccgttccccg	cccgcggacc	gaccagactt	8100
cggccagctg	cggcttcagt	tcttgacca	gctctgcag	cgtttcgtcg	ctgggccaat	8160
cgtgaatacg	gaagttgtcg	ggtagaggcg	ggaggcggcg	gacttccagg	aggtgtgtga	8220
gggcccgcag	gagatgcagg	tggctactga	tttcccacgg	atgacggtcg	cgggcgtcca	8280
aggcgaagag	atgaccgtgg	ggccgcggcg	ccaccagcgt	tccgcggggg	gtctttatcg	8340
gcggcgggga	cgggtccccg	gcggcagcgg	cggctcggga	cccgcgggca	agtcgggcag	8400
cggcacgtcg	gcgtggagct	cgggcagggg	ctggtgctgc	gcgcggagct	gactggcaaa	8460
ggctatcacc	cggcgattga	cgtctctggat	ccggcggcgc	tgctgaaga	ccaccggacc	8520
cgtggtcttg	aacctgaaag	agagttcgac	agaatcaatc	tccgcatcgt	taaccgcggc	8580
ctggcgcagg	atttcggcca	cgtccccgga	gttgtcttga	tacgcgattt	ctgcatgaa	8640
ctggtcgatt	tcctcttcct	gcaagtctcc	gtgaccggcg	cgttcgacgg	tggccgcgag	8700
atcgttgag	atgcggccca	tgagctggga	aaaggcattg	atgccgacct	cgttccacac	8760
tcggctgtac	accacctctc	cgtgaacgtc	gcgggcgcgc	atcaccacct	gggcgagatt	8820
gagttccacg	tggcgggcga	aaaccggata	gtttcggagg	cgctgataca	gatagttgag	8880
ggcgtggcg	gcgtgctcgg	ccacaaaaaa	atacatgac	cagcggcgga	gggtcagctc	8940
gttgatgtcg	cccagcgctt	ccaggcgttc	catggcctcg	taaaagtcca	cggcaaagtt	9000
gaaaaattgg	ctgttcctgg	ccgagaccgt	gagctcttct	tccaagagcc	gaatgagatc	9060
cggcacggcg	gccctgactt	cgcgttcgaa	agccccgggt	gcctcctcca	cctcttcctc	9120
ctcgacttct	tcgaccgctt	cgggcacctc	ctcttcctcg	accaccacct	caggcggggc	9180
tcggcggcgc	cggcggcgga	cgggcaggcg	gtcgacgaaa	cgtcgatca	tttccccctt	9240
ccgtcgacgc	atggtctcgg	tgacggcgcg	accctgttcg	cgaggacgca	gggtgaaggc	9300
gccgcgcgcg	agcggaggta	acaggggagat	cggggggcgg	tcgtggggga	gactgacggc	9360
gctaactatg	catctgatca	atgtttgcgt	agtgacctcg	ggtcggagcg	agctcagcgc	9420
ttgaaaatcc	acgggatcgg	aaaaccgttc	caggaaacgcg	tctagccaat	cacagtcgca	9480
aggtaagctg	aggaccgtct	cgggggcttg	tctgttctgt	cttcccgcgg	tggtgctgct	9540
gatgaggtag	ttgaagtagg	cgctcttgag	gcggcggatg	gtggacagga	gaaccacgtc	9600
tttgcccca	gcttgctgta	tccgcaggcg	gtcggccatg	ccccacactt	ctccttgaca	9660

gcggcgagg	tccttgtagt	attcttgcat	cagcctttcc	acgggcacct	cgtcttcttc	9720
ttccgctcgg	ccggacgaga	gccgcgtcag	gccgtacccg	cgtgccccct	gtggttgagg	9780
cagggccagg	tcggccacga	cgcgctcggc	cagcacggcc	tgctggatgc	gggtgagggt	9840
gtcctgaaag	tcgtcgagat	ccacaaagcg	gtggtacgcg	ccagtgttga	tggtgtaggt	9900
gcagttgctc	atgacggacc	agtttacggg	ctgggtgcc	tggtccacgg	tttccaggta	9960
gcggagacgc	gagtagggcc	gcgtctcgaa	gatgtagtcg	ttgcagggtcc	gcagcaggta	10020
ctggtagccc	accagcagat	gcggcgggcg	ctggcggtag	aggggccacc	gctgggtggc	10080
gggggcgttg	ggggcgagat	cttccaacat	gaggcggtga	tagccgtaga	tgtagcgcga	10140
catccaagtg	atgccgctgg	ccgtgggtgct	ggcgcgggcg	tagtcgcgaa	cgcggttcca	10200
gatgtttcgc	agcggctgga	agtactcgat	ggtggggcga	ctctgccccg	tgaggcgggc	10260
gcagtcggcg	atgctctacg	gggaaaaaga	agggccagtg	aacaaccgcc	ttccgtagcc	10320
ggaggagAAC	gcaagggggg	caaagaccac	cagggctcgg	gttcgaaacc	cgggtggcgg	10380
cccgaatacg	gagggcggtt	ttttgctttt	ttctcagatg	catcccgtgc	tgcggcagat	10440
gcgtccgaac	gcgggggtccc	agtcccgggc	ggtgcctgcg	gccgtgacgg	cggcttctac	10500
ggccacgtcg	cgtccacccc	cgcctaccac	ggcccaggcg	gcggtggctc	tgcgcggcgc	10560
aggggaaccc	gaagcagagg	cgggtgttga	cgtggaggag	ggccaggggt	tggtcggct	10620
gggggccctg	agtcccagac	ggcaccgcg	cgtggctctg	aagcgcgacg	cggcggaggc	10680
gtacgtgccg	cggagcaatc	tgtttcgcga	ccgcagcggc	gaggaggccg	aggagatgcg	10740
agacttgctg	tttcggggcg	ggaggaggtt	gcgtcacggg	ctggaccggc	agagggttct	10800
gagagaggag	gactttgagg	cggacgagcg	cacgggggtg	agtcccgcgc	gggctcacgt	10860
ggcgcccgcc	aacctggtga	gcgcgtacga	gcagacggtc	aaggaggaga	tgaacttcca	10920
gaagagcttc	aatcatcacg	tgccgacgct	gattgcgcgc	gaagagggtg	ccatcggcct	10980
catgcatctg	tgggattttg	tggaggcgta	cgttcagaac	cccagcagca	agccgctgac	11040
ggctcagctg	ttcctcatcg	tgcaacatag	tcgagacaac	gaaacgttca	gggaggccat	11100
gctgaacatt	gcagagcctg	aggggcgctg	gctcttggtg	ctcattaaca	tcttgccagag	11160
tatcgtagtg	caggagcgct	cgtgagcct	ggccgacaag	gtggctgcc	tcaactacag	11220
catgctgtcg	ctgggcaa	tttacgccc	caagatctac	aagtctccgt	tcgtcccat	11280
agacaaggag	gtgaagatag	acagctttta	catgcgcgtg	gcgtcaagg	tgctgactct	11340
aagcgacgac	ctgggggtgt	accgcaacga	ccgcatacac	aaggcggtga	gcgccagccg	11400

ccggcgcgag ctgagcgacc gcgagctttt gcacagcctg catcggggcgt tgactgggtgc 11460
 cggcagcgcc gagggggccg agtactttga cgccggagcg gacttgcgct ggcagccatc 11520
 ccgacgcgcg ctggaggcgg ctggcgctcg ggagtacggg gtcgaggacg acgatgaagc 11580
 ggacgacgag ttgggcattg acttgtagcc gtttttcggt agatatgtcg gcgaacgagc 11640
 cgtctgcggc cgccatgggt acggcgggcg gcgcgcccc aagcccgcc acgcgcgcg 11700
 cgctgcagag tcagccttcc ggagtgcgc ccgcggacga ctgggtccgag gccatgcgtc 11760
 gcacctggc gctgacggcg cgcaaccccg aggtttttcg gcagcagccg caggcaaacc 11820
 ggtttgcgcc cattttggaa gcggtgggtgc cctccagacc caacccacc caggaagc 11880
 tgctggccat cgtcaacgcc ctggcgga ccaaggccat ccgcccagac gaggccggcg 11940
 aggtttacaa cgcgctgcta gaaagggtgg gacgctacaa cagctccaac gtgcagacca 12000
 atctggaccg cttgggtgac gacgtgaagg aggcgtagc ccagcgagag cggtttttca 12060
 aggaagccaa tctgggctcg ctgggtggccc tcaacgcctt cctgagcacg ctgcccggca 12120
 acgtgccccg gggtcaggag gactacgtga actttctgag cgccctccgc ctgatgggtg 12180
 ccgaggtgcc gcagagcgag gtgtaccagt ctggcccaa ctactacttc cagacctccc 12240
 ggcagggcct gcagacggta aacctgacgc aggcctttca gaacctgcag ggcctttggg 12300
 ggggtgcgcg tccgctgggc gaccgcagca cgggtgtccag cctgctgacc cccaatgccc 12360
 ggctgctctt gcttctcatt gctccgttca ccgacagcg ttccatcagc cgcgactctt 12420
 acctgggaca cctgctcacc ctgtaccggg aggcctcagg gcaggcgcg gtggacgagc 12480
 agacgtacca ggaaatcacc agcgtgagcc gcgcgctggg gcaggaggac acgggcagct 12540
 tggaggcgac tctgaacttc ctgctgacca accggcgga gcgcctacct cccagtagc 12600
 cgctgaacgc ggaggaggag cgcctcctgc gtttcgtgca gcagagcacc gcgctgtact 12660
 tgatgcggga aggcgcctct cccagcgctt cgctggacat gacggcgcc aacatggagc 12720
 catcgttcta cgccgcaac cgtcccttcg tcaaccggct aatggactat ttgcatcggg 12780
 cggcgccct gaaccggaa tactttacta acgtcatcct gaacgaccgt tggtgccac 12840
 ctcccgctt ctacacggg gagttcgacc tcccgaggc caacgacggt ttcattgtgg 12900
 acgacgtgga cagcgtgttc ctgcccggca agaaggaggc gggtgactct cagagccacc 12960
 gcgcgagcct cgagacctg gggcgaccg ggccgcgtc tccgctgcct cgcctgccga 13020
 gcgcagcag cgccagcgtg gggcggtga gccgtccgc cctcagcggg gaggaggact 13080

ggtggaacga tccgctgctc cgtccggccc gcaacaaaaa ctcccccaac aacgggatag 13140
 aggatttggg agacaaaatg aaccgttggg agacgtatgc ccaggagcat cgggagtgge 13200
 aggcgaggca acccatgggc cctgttctgc cgccctctcg gcgcccgcgc agggacgaag 13260
 acgccgacga ttcagccgat gacagcagcg tgttgatctt gggcgggagc gggaaaccct 13320
 ttgccacact gcaacctcgc ggcgtgggtc ggcggtggcg ctaggaaaaa aaattattaa 13380
 aagcacttac cagagccatg gtaagaagag caacaaaggt gtgtcctgct ttcttcccg 13440
 tagcaaatg cgtcgggcgg tggcagttcc ctccgcggca atggcgtag gcccgcctcc 13500
 ttcttacgaa agcgtgatgg cagcggccac cctgcaagcg ccgttgagga atccttacgt 13560
 gccgcgcga tacctggagc ctacgggcgg gagaaacagc attcgttact cggagctgac 13620
 gccctgtac gacaccaccc gcctgtacct ggtggacaac aagtcagcag atatcgccac 13680
 cttgaactac cagaacgacc acagcaactt tctcacgtcc gtggtgcaga acagcgacta 13740
 cacgcccgc gaagcgagca cgcagaccat taacttgagc gaccgctcgc gctggggcgg 13800
 ggacttgaaa accattctgc acactaacat gcccaacgtg aacgagttca tgtttaccaa 13860
 ctcgttcagg gctaaactta tgggggcgca cgaggccgac aaggaccggg tttatgagtg 13920
 ggtgcagctg acgctgccgg aggggaactt ttcagagatt atgaccatag acctgatgaa 13980
 caacgccatt atcgaccact acctggcggg agccagacag caggggggta aagaaagcga 14040
 gatcggcgtc aagtttgaca cgcgcaactt tcgtctgggc tgggaccggg agacggggct 14100
 tgtgatgccg ggggtgtaca cgaacgaagc ttccatccc gacgtggctc tcttgccggg 14160
 ctgcggggtg gactttacct acagccggtt aaacaacctg ctaggcatac gcaagagaat 14220
 gccctttcag gaagggtttc agatcctgta cgaggacctg gagggcggta acatcccggc 14280
 cctgctggac gtgccggcgt acgaggagag catcgccaac gcaaggagg cggcgatcag 14340
 gggcgataat ttccgggcgc agcccaggc ggctccaacc ataaaaccg ttttggaaga 14400
 ctccaaaggg cggagctaca acgtaatagc caacaccaac aacacggctt acaggagctg 14460
 gtatctggct tataactacg gcgacccgga gaagggggtt agggcctgga ccctgctcac 14520
 cactccggac gtgacgtgcg gttcagagca ggtctactgg tcgctgcctg acatgtacgt 14580
 ggaccctgtg acgtttcgct ccacgcagca agttagcaac taccagtggt tgggagcgga 14640
 gcttatgccg attcacagca agagctttta caacgagcag gccgtctact cacagctcat 14700
 tcgtcagacc accgccctaa cgcacgtttt caaccgcttc cccgagaacc aaatcctagt 14760
 gcgacctcca gcgcccacca tcaccacgt cagcgagaac gtgcccgtc taaccgatca 14820

cgggacgctg	cctttgcaga	acagcatccg	cggagttcag	cgagttacca	tacaggacgc	14880
ccgtcgtcgg	acctgtccct	acgtctacaa	agccttggga	atcgtggccc	cgcgctcct	14940
gtcgagtcgc	actttctaga	tgtccatcct	catctctccc	agcaacaata	ccggttgggg	15000
tctgggctg	acaaaaatgt	acggaggcgc	caaacgacgg	tcccacaac	atcccgtg	15060
agtgcgctgg	cacttttagag	ccccatgggg	gtcgcacacg	cgcgggcgca	ccggccgaac	15120
caccgtcgac	gacgtgatcg	atagcgtggg	ggccgacgcc	cgcaactacc	agcccgtc	15180
atccacgggtg	gacgaagtca	tcgacggcgt	ggtagccgac	gccagggcct	acggccgcag	15240
aaagtctcgt	ctgcgccgcc	gccgttcgct	aaagcgcccc	acggccgcca	tgaagccgc	15300
tcgtctctcg	ctgcgtcgcg	cacgtatcgt	gggtcgccgc	gccgccagac	gcgcagccgc	15360
caacgccgcc	gccggccgag	tgcgccgccg	ggccgccag	caggccgccg	cgccatctc	15420
cagtctatcc	gcccccgac	gcgggaatgt	gtactgggtc	agggactcgg	ccaccggcgt	15480
gcgagttccc	gtgagaaccc	gtcctcctcg	tccctgaata	aaaagttcta	agcccaatcg	15540
gtgttcctgt	gtgtgttcag	ctcgtcatga	ccaaacgcaa	gtttaagag	gagctgctgc	15600
aagcgtggt	ccccgaaatc	tatgcgccgg	cgccggacgt	gaaaccgcgt	cgcgtaaac	15660
gcgtgaagaa	gcaggaaaag	ctagagacaa	aagaggaggc	ggtggcggtg	ggagacgggg	15720
aggtggagtt	tgtgcgtcgc	ttcgtgccgc	gtcggcgagt	gaattggaag	gggcgcaagg	15780
tgcaacgggt	gctgcgtccc	ggcacggtgg	tgtctttcac	cccgggtgaa	aaatccgcct	15840
ggaaggccat	aaagcgctg	tacgatgagg	tgtacgggga	cgaagacatt	ctggagcagg	15900
cgctggatag	aagcggggag	tttgcttacg	gcaagagggc	gaggacgggc	gagatcgcca	15960
tcccgtgga	cacttccaac	cccaccccca	gtctgaaacc	cgtgacgctg	caacaggtgt	16020
tgccggtgag	cgccccctcg	cgacgcggca	taaaacgcga	gggcggcgag	ctgcagccca	16080
ccatgcagct	cctgggtccc	aagaggcaga	aactagagga	cgtactggac	atgataaaaa	16140
tgagagccga	cgtgcagccc	gatattaaaa	tccgtcccat	caaagaagtg	gcgccgggaa	16200
tgggcggtgca	gaccgtggac	atccagattc	ccatgaccag	cgccgcacag	gcggtagagg	16260
ccatgcagac	cgacgtgggg	atgatgacgg	acctgcccgc	agctgctgcc	gccgtggcca	16320
gcgccgcgac	gcaaacggaa	gccggcatgc	agaccgaccc	gtggacggag	gcgcccgtgc	16380
agccggccag	aagacgcgtc	agacggacgt	acggccccgt	ttctggcata	atgccggagt	16440
acgcgctgca	tccttccatc	atccccaccc	ccggtaccg	ggggcgacac	taccgtccgc	16500

gacgcagcac cactcgccgc cgtcgcccga cggcacgagt cgccaccgcc agagtgcagac 16560
 gcgtaacgac acgtcgccgc cgccgcttga ccctgcccggt ggtgcgctac catcccagca 16620
 ttcttttaaaa aaccgctcct acgttgccaga tgggcaagct tacttgctga ctccgtatgg 16680
 ccgtgcccgg ctaccgagga agatcccgcc gacgacggac tttgggaggg agcgggtttgc 16740
 gccgcccgtcg ggcgggttcac cggcgccctca agggaggcat tctgccggcc ctgatcccca 16800
 taatcgccgc agccatcggg gccattcccg gaatcgccag cgtagcggtg caggctagcc 16860
 agcgccactg attttactaa ccctgtcggg cgcgccgtct ctttcggcag actcaacgcc 16920
 cagcatggaa gacatcaatt tctcctctct ggcccccggg cagggcacgc ggcggtatat 16980
 ggggacgtgg agcgagatcg gcacgaacca gatgaacggg ggcgctttca attggagcgg 17040
 tgtgtggagc ggcttgaaaa atttcgggtc cactctgaaa acttacggca accgggtgtg 17100
 gaactccagc acggggcgaga tgctgaggga caagctaaag gacacgcagt ttcagcaaaa 17160
 ggtggtggac ggcacgctt cgggcctcaa cggcgccgct gacctggcca accaggccat 17220
 tcaaaaggaa attaacagcc gcctggagcc gcgcccgag gtggaggaga acctgcccc 17280
 tctggaggcg ctgcccccca agggagagaa gcgcccgcgg cccgacatgg aggagacgct 17340
 agttactaag agcgaggagc cgccatcata cgaggaggcg gtgggtagct cgcagctgcc 17400
 gtccctcagc ctgaagccca ccacctatcc catgaccaag cccatcgctt ccatggcgcg 17460
 ccccggtggga gtcgaccgc ccacgacgc ggtggccact ttggacctgc cgcgccccga 17520
 accgggcaac cgcgtgcctc ccgtcccat cgctccgccc gtttctcgcc ccgcatccg 17580
 ccccgctgcc gtggccactc cccgctatcc gagccgcaac gccaaactggc agaccaccct 17640
 caacagtatt gtcggactgg gggtgaaagtc tctgaagcgc cgtcgctgtt tttaaagcac 17700
 aatttattaa acgagtagcc ctgtcttaac ccacgcttgt atgtgtgctt atatcagcg 17760
 ttcagagcct gaccgtccgt caagatggcc actccgtcga tgatgccgca gtggtcgtac 17820
 atgcacatcg ccgggcagga cgctcggag tacctgagcc cgggtctggt gcagtttgcc 17880
 cgtgcgacgg aaacctactt ctactgggc aacaagttca ggaacccac cgtggcgccc 17940
 acccacgacg tcaccaccga tcggtccag cgactgacaa tccgcttcgt ccccggtggac 18000
 aaggaagaca ccgcttactc ctacaaaacc cgcttcacgc tggccgtggg cgacaaccgg 18060
 gtgctagaca tggccagtac ctactttgac atccgcgggc tgatcgaccg cggacctagc 18120
 ttcaagcctt actccggcac ggcttacaac tctactggct ccaaagggg gcccaacaac 18180
 agccaatgga acgccacaga taacgggaac aagccagtgt gttttgctca ggcagctttt 18240

ataggtcaaa gcattacaaa agacggagtg caaatacaga actcagaaaa tcaacaggct 18300
gctgccgaca aaacttacca accagagcct caaattggag ttccacctg ggataccaac 18360
gttaccagta acgctgccgg acgagtgtta aaagccacca ctcccatgct gccatgttac 18420
ggttcatatg ccaatccac taatccaaac gggggtcagg caaaaacaga aggagacatt 18480
tcgctaaact ttttcacaac aactgcggca gcagacaata atcccaaagt ggttctttac 18540
agcgaagatg taaaccttca agccccgat actcacttag tatataagcc aacgggtggga 18600
gaaaacgtta tcgccgcaga agccctgcta acgcagcagg cgtgtcccaa cagagcaaac 18660
tacataggtt tccgagataa ctttatcggg ttaatgtatt ataacagcac agggaacatg 18720
ggagtctcgg caggtcaggg ctgcaggtta aacgcagttg tagacctgca agatcgaaac 18780
acggaactgt cctatcagct aatgctagat gctctgggtg acagaactcg atatttctca 18840
atgtggaatc aggccgtgga cagctacgat ccagacgtta ggattatcga gaacctggg 18900
gtggaagacg agctgccc aa ttactgtttt ccaactccag gcatgggtat ttttaactcc 18960
tacaaggggg taaaaccaca aaatggcggg aatggtaact gggaagcaaa cggggaccta 19020
tcaaatgcc aatgagatgc tttaggaaac atttttgcc tggaaattaa cctccacgca 19080
aacctgtggc gcagcttctt gtacagcaat gtggcgctgt acctgccaga cagctataaa 19140
ttcaactccg ctaacatcac tctgcccgcc aacaaaaaca cctacgagta tatcaacggg 19200
cgcgtcactt ctccaacctt ggtggacacc tttgttaaca ttggagcccg atggtcgcgg 19260
gatcccatgg acaacgtcaa cccctttaac catcaccgga acgcgggcct ccgttaccgc 19320
tccatgctgc tgggaaatgg acgcgtgggt cctttccaca tacaagtgc gcaaaatatt 19380
ttcgcgatta agaacctcct gcttttgccc ggctcctaca cttacgagtg gagcttcaga 19440
aaagacgtga acatgattct gcagagcacc ctgggcaatg atpttcgagt ggacggggcc 19500
agcgtccgca ttgacagcgt caacttgtag gccaaacttt tccccatggc gcacaacacc 19560
gctttctact tgggaagccat gctgcgaaac gacaccaacg accagtcggt taacgactac 19620
ctcagcgagg ccaacatgct ttatccatt cgggccaacg ccaccaacgt tccatttcc 19680
attccctccc gcaactgggc ggccttcggg ggatggagct tcaccgcct taaagccaag 19740
gaaacgcctt ccttgggctc cggctttgac ccctactttg tgtactcagg caccattcct 19800
tacctggacg gcagctttta cctcaaccac actttcaaac gtctgtccat catgttcgat 19860
tcttccgtaa gttggccggg caacgacgc ctctgacgc cgaacgagtt cgaaattaag 19920

cgcatgtgtg	acggggaagg	ctacaacgtg	gctcaaagta	acatgaccaa	agaactggttt	19980
ttaattcaaa	tgctcagcca	ctacaacatc	ggctaccaag	gcttctatgt	ccccgagggc	20040
tacaaggatc	ggatgtatc	tttcttccga	aactttcagc	ccatgagccg	ccaggtgccg	20100
gatcccaccg	ctgccggcta	tcaagcgtt	cccctgccc	gacaacacaa	caactcgggc	20160
tttgtggggt	acatgggccc	gaccatgcgc	gaaggacagc	cataccggc	caactacccc	20220
tatcccctga	tggcgctac	cgccgtcccc	gccattacc	agaaaaagtt	tttgtgcgac	20280
cgcgctcatgt	ggcgcatacc	tttttcagc	aactttatgt	caatgggggc	cctgaccgac	20340
ctcggacaga	acatgcttta	cgctaactcc	gcccattgcc	tggatatgac	ttttgaggtg	20400
gaccccatga	acgagcccac	gttgctgtac	atgctttttg	aggtgttcga	cgtggtcaga	20460
gtgcaccagc	cgcaccggcg	tattatcgag	gccgtgtacc	tgcgcacccc	cttctctgcg	20520
ggcaatgcca	ccacataagc	cgctgaacta	gctgggtttt	acccagatc	ccatgggctc	20580
cacggaagac	gaactgcggg	ccattgtgcg	agacctgggc	tgcggaccct	acttctggg	20640
cacctttgac	aagcgtttc	cggggttcgt	gtctcctcgc	aaactcgcgt	gcgcgatcgt	20700
gaataccgcc	ggcgcagaga	cgggaggaga	gcattggcta	gctctgggct	ggaacccccg	20760
ctcgtccacg	tttttctgt	tgcacccctt	tggtttttca	gaccaacgct	tgaagcagat	20820
ctatgcattt	gaatatgagg	gtctactcaa	gcgaagcgcg	ctggcctcct	ccgccgatca	20880
ctgtctaacc	ctggtaaaga	gcactcagac	ggttcagggc	cctcacagcg	ccgcctgtgg	20940
ccttttttgt	tgcatgtttt	tgcacgcctt	tgtgaactgg	ccggacaccc	ccatggaaaa	21000
caacccccacc	atggacctcc	tgactggcgt	tcccaactcc	atgctccaaa	gcccagcgt	21060
gcagaccacc	ctcctccaaa	accagaaaaa	tctgtacgcc	tttctgcaca	agcactctcc	21120
ctactttcgc	cgccatcggg	aacaatatga	aaatgcaacc	gcgtttaaca	aaactctgta	21180
acgtttaata	aatgaacttt	ttattgaact	ggaaaacggg	tttgtgattt	ttaaaaatca	21240
aaggggttga	gctggacatc	catgtgggag	gccggaaggg	tggtgttctt	gtactggtac	21300
ttgggcagcc	acttaaactc	tggaatcaca	aacttgggca	gcggtatttc	tggaagttg	21360
tcgtgccaca	gctggcgggt	cagctgaagt	gcctgcagaa	catcgggggc	ggagatcttg	21420
aagtcgcagt	ttatctgggt	cacggcacgc	gcgttgcggt	acatgggatt	ggcacactga	21480
aacaccagca	ggctgggatt	cttgatgcta	gccagggcca	cggcgtcggg	cacgtcaccg	21540
gtgtcttcta	tggtggacag	cgaaaaaggc	gtgactttgc	aaagctggcg	tcccgcgcga	21600
ggcacgcaat	ctcccaggta	gttgactca	cagcggatgg	gcagaagaag	atgcttgtgg	21660

ccgcgggtca	tgtagggata	ggccgctgcc	ataaaaagctt	cgatctgcct	gaaagcctgc	21720
ttggccttgt	gcccttcggt	ataaaaaaca	ccgcaggact	tggtggaaaa	ggtattactg	21780
gcgcaagcgg	catcgtgaaa	gcaagcgcg	gcgtcttcgt	ttcgtaactg	caccacgctg	21840
cggccccacc	ggttctgaat	caccttggcc	ctgccgggg	tttccttgag	agcgcgctgg	21900
cgggcttcgc	tgcccacatc	catttcacag	acatgctcct	tgtaatcat	ggccagaccg	21960
tggaggcagc	gcagctcctc	gtcatcgctc	gtgcagtgat	gctcccacac	gacgcagcca	22020
gtgggctccc	acttgggctt	ggaggcctcg	gcaatgccag	aatacaggag	aacgtagtgg	22080
tgcagaaaac	gtcccatcat	ggtgccaaag	gtttcttggc	tgctgaaggt	catcgggcag	22140
tacctccagt	cctcggttaag	ccaagtgttg	cagatcttcc	tgaagaccgt	gtactgatcg	22200
ggcataaagt	ggaactcatt	gcgctcggtc	ttgtcgatct	tatacttttc	catcagacta	22260
tgcataatct	ccatgccctt	ttcccaggcg	caaacaatct	tggtgctaca	cgggtaggt	22320
atggccaaag	tggttggcct	ctgaggcggc	gcttgttctt	cctcttgagc	cctctcccg	22380
ctgacggggg	ttgaaagagg	gtgccccttg	gggaacggct	tgaacacggg	ctggcccag	22440
gcgtcccgaa	gaatctgcat	cgggggattg	ctggcgcgtc	tggcgatgat	ctgaccccg	22500
ggctcctcca	cttcgtcctc	ctcgggaactt	tcctcgctgt	tttcggggga	cggtagggga	22560
gtagggggaa	gagcgcgggc	cgccttcttc	ttgggcggca	gttcgggagc	ctgctcttga	22620
cgactggcca	ttgtcttctc	ctaggcaaga	aaaacaagat	ggaagactct	ttctcctcct	22680
cctcgtcaac	gtcagaaagc	gagtcctcca	ccttaagcgc	cgagaactcc	cagcgcatag	22740
aatccgatgt	gggctacgag	actccccccg	cgaacttttc	gccgcccccc	ataaacacta	22800
acgggtggac	ggactacctg	gccctaggag	acgtactgct	gaagcacatc	aggcggcaga	22860
gcgttatcgt	gcaagatgct	ctcacgcgagc	gactcgcggg	tccgctggaa	gtggcggaac	22920
ttagcgccgc	ctacgagcga	accctcttct	ccccaaagac	tcccccaag	aggcaggcta	22980
acggcacctg	cgagcctaac	cctcgactca	acttctacce	tgcccttgcc	gtgccagagg	23040
tactggctac	gtaccacatt	tttttccaaa	accacaaaat	ccctctctcg	tgccgcgcca	23100
accgcaccaa	agccgatcgc	gtgctgcgac	tggaggaagg	ggctcgcata	cctgagattg	23160
cgtgtctgga	ggaagtccca	aaaatctttg	aaggctctgg	ccgcgacgaa	aagcgagcag	23220
caaacgctct	ggaagagAAC	gcagagagtc	acaacagcgc	cttggtagaa	ctcgagggcg	23280
acaacgccag	actggccgtc	ctcaaacggg	ccatagaagt	cacgcacttc	gcctaccccg	23340

ccgttaacct ccctccaaaa gttatgacag cggatcatgga ctcgctgctc ataaagcgcg 23400
 ctcagccctt agaccagag cagcaaaaaca acagtacga aggaaaaccg gtgggttctg 23460
 atgaggagtt gagcaagtgg ctgtcctcca acgacccgc cacgttgag gaacgaagaa 23520
 aaaccatgat ggccgtggtg ctagttaccg tgcaattaga atgtctgcag aggttctttt 23580
 cccaccaga gaccctgaga aaagtggagg aaacgctgca ctacacattt aggcacggct 23640
 acgtgaagca agcctgcaag atttccaacg tagaacttag caacctcctc tcctacctgg 23700
 ggatcttgca cgaaaaccgc ctcggacaaa acgtgctgca cagcactg aaaggagaag 23760
 cccgcccaga ctatgtgcga gactgctgt tcctagcgt agtgtacacc tggcagagcg 23820
 gaatgggagt ctggcagcag tgcctggagg acgaaaacct caaagagctt gaaaagctgc 23880
 tggcgctc cagaaggga ctgtggacca gttttgacga gcgcaccgc gcgcgagacc 23940
 tagctgat ttttttctt cccaagctgg tgcagactct ccgggaagga ctgccagatt 24000
 ttatgagtc aagcatcttg caaaacttcc gctcttctat cttggaacgc tcgggaatct 24060
 tgcccgcac tagctgcgc ctaccacag attttgtgct tctccactac cgcgaatgcc 24120
 caccgcccgt gtggccgtac acttacttgc ttaaactggc caactttcta atgttccact 24180
 ctgacctggc agaagacgtt agcggcgagg ggctgctaga atgccactgc cgctgcaacc 24240
 tgtgcacccc ccacgctct ctagtatgca aactccccct gctcaatgag acccagatca 24300
 tcggtacctt tgaaatccag ggaccctccg acgcgaaaaa cggcaagcag gggctctggc 24360
 taaaactcac agccggactg tggacctccg cctacttgcg caaatttgta ccagaagact 24420
 atcacgccc ccaaatataa ttttacgaaa accaatcaaa accacccaaa agcagagtaa 24480
 cggcttgct cattacgcag agcagcatag ttgggcagtt gcaagccatt aacaaagcg 24540
 ggcaagagtt tctcctaaaa aaaggaaaag gggcttactt ggacccccag accggcgagg 24600
 aactcaacgg accctcctca gtgcgaggtt gtgtgcccc tgccgcccc aaagaacacc 24660
 tcgcagtga acatgccaga gacggaggaa gaggagtga gcagtgtgag caacagcgaa 24720
 acggaggaag agcgtggc cgaggggtgc aacggggaag aggacacgga gggacggcga 24780
 agtcttcgcc gaagaactct cgcgctgcc cccgaagtcc cagccggccg cctcggccca 24840
 agatcccga cacaccgta gatgggatag caagacaaa aagccgggta agagaaacgc 24900
 tcgccccgc cagggtacc gtcgtggag aaagcacaaa aactgcatct tatcgtgctt 24960
 gctccagtgc ggcgagacg ttctgttcac ccgtagatag ttgcttttta acaaaggggt 25020
 ggccgtcccc cgtaacgtcc tccactacta ccgtcactct tacagctccg aagcgagcg 25080

ctaagaaaac gcagcagttg ccggcgggag gactgcgtct cagcgccccga gaacccccag 25140
 ccaccagggga gctccgaaac cgcataattc ccaccctcta cgctatcttt cagcaaagcc 25200
 gggggcagca gcaagaactg aaaataaaaa accgcacgct gaggtcgctt acccgaagct 25260
 gcctctatca caagagcgaa gagcagctgc agcgaaccct ggaggacgca gaagcgctgt 25320
 tccagaagta ctgcgcgacc accctaaata actaaaaaag cccgcgcgcg ggacttcaaa 25380
 ccgtctgacg tcaccagccg cgcgcacaaa tgagcaaaga gattcccacg ccttacatgt 25440
 ggagttacca gccgcagatg ggattagccg ccggcgcgcg ccaggattac tccacgaaaa 25500
 tgaactggct cagcgccggg ccccatga tttcccgct aaacgacatt cgcgccacc 25560
 gcaatcagct attgttagaa caggctgctc tgaccgccac gccccgaat aacctgaacc 25620
 ctcccagctg gccagctgcc ctggtgtacc aggaacgcc tccaccacc agcgtacttt 25680
 tgccccgtga cgcaccggcg gaagtccaga tgactaacgc gggcgcgcaa ttagcgggcg 25740
 gatcccggtt tcggtacaga gttcacggcg ccgcacccta tagcccaggt ataaagaggc 25800
 tgatcattcg aggcagaggt gtccagctca acgacgagac agtgagctct tcgcttggtc 25860
 tacgaccaga cggagtgttc cagctcgcgg gctcgggccc ctcttcgttc acgcctcgcc 25920
 aggcatacct gactctgcag agctctgcct ctccagctcg ctccggaggga atcggacccc 25980
 ttcagtttgt ggaggagttt gtgcctcgg tctactttca gcctttctcc ggatcgcccg 26040
 gccagtacc gccagagttc atcccdaact tcgacgcggg gagtgactct gtggacgggt 26100
 atgactgatg tcgagcccg ctcagtgtta gtggaacaag cgcggctcaa tcacctgggt 26160
 cgttgccgcc gccgctgctg cgtggctcgc gacttgagct tagctctcaa gtttgtaaaa 26220
 aaccgctccg aaaccgggag cgtgtgcac ggggtggagc tagtgggtcc tgagaaggcc 26280
 accatccacg ttctcagaaa ctttgtggaa aaaccattt tgggtaaacg agatcagggg 26340
 ccttttgtaa tcagcttact ctgcacctgt aacctgttg accttcacga ctattttatg 26400
 gatcatttgt gcgctgaatt caataagtaa agcgaattct taccaagatt atgatgtcca 26460
 tgactgttcc tcgccactat acgatgttgt gccagtaaac tctcttgctg acatctatct 26520
 gaactgttcc ttttgggtccg cacagcttac ttggtactac ggtgacaccg tcctttctgg 26580
 ctactgggc agctcacacg gaataaact tcacctcttt tcgccgtttc gatacggaaa 26640
 ctacagctgt cgtgccggta cctgcctcca cgttttcaat ctccagccct gtccaccgac 26700
 caaacttgta tttgtcgact ctaagcactt acagctcaac tgcagcatte taggccccag 26760

tatcttgtgg	acatacaata	aaatcagggt	ggtggaattt	gtctactacc	caccagcgc	26820
ccgcggtttt	ggggaaatc	ctttccagat	ctactacaac	tatcttgcca	cacattatgc	26880
aagtcaacag	caactaaact	tgcaagcacc	cttcacgcca	ggagagtact	cctgtcacgt	26940
aggctcctgc	acagaaactt	ttattctctt	caacagatct	tctgccattg	aacgcttcac	27000
tactaactac	tttagaaacc	aagttgtgct	tttactgac	gaaacccta	acgtcaccct	27060
ggactgtgca	tgtttttctc	atgacaccgt	aacttggaact	cttaacaata	ctctctggct	27120
cgcgttcgat	aaccaaagct	tgattgttaa	aaattttgat	ttaaccttta	ctaaaccctc	27180
tectcgcgaa	atagttatct	ttgctccttt	taatccaaa	actaccttag	cctgtcaggt	27240
tttgtttaag	ccttgccaaa	caaacttta	gtttgtttat	ttgctccgc	aatctgtcaa	27300
actcatagaa	aaatacaaca	aagcgccgt	cttggtcct	aaaaccttct	accactggct	27360
aacctacacg	gggctgtttg	cactaattgt	tttttcccta	attaacattt	ttatatgttt	27420
cttgccctcc	tccttctttt	cgcgaacacc	gttgccgcag	aaagacctct	ccttattact	27480
gtagcgcttg	ctatacaaaa	ccaagagtgg	tcaaccgtgc	tctcaatcta	ttttcaattt	27540
ttcattttgt	ccttaatact	ttctcttatt	gtcgtaaca	atgatctgga	gcattggctct	27600
cgcctttttt	tggttgctta	gtgcaaaagc	cactattttt	cacaggtatg	tggaagaagg	27660
aactagcacc	ctctttacga	tacctgaaac	aattaaggcg	gctgatgaag	tttcttggta	27720
caaaggctcg	ctctcagacg	gcaaccactc	attctcagga	cagacccttt	gcacccaaga	27780
aacttatttt	aaatcagaac	tacaatacag	ctgcataaaa	aactttttcc	atctctacaa	27840
catctcaaaa	ccctatgagg	gtatttacaa	tgccaagggt	tcagacaact	ccagcacacg	27900
gaactttttac	tttaatctga	cagttattaa	agcaatttcc	attcctatct	gtgagttag	27960
ctcccagttt	ctttctgaaa	cctactgttt	aattactata	aactgcacta	aaaatcgct	28020
tcacaccacc	ataatctaca	atcacacaca	atcaccttgg	gttttaaac	taaaattttc	28080
tccacacatg	ccttcgcaat	ttctcagca	agttaccgtc	tctaacataa	gcaagcagtt	28140
tggtttttac	tatcctttcc	acgaaetgtg	cgaaataatt	gaagccgaat	atgaaccaga	28200
ctactttact	tacattgcca	ttggtgtaat	cgttgtttgc	ctttgctttg	ttattggggg	28260
gtgtgtttat	ttgtacattc	agagaaaaat	attgctctcg	ctgtgctcct	gcggttacaa	28320
agcagaagaa	agaattaaaa	tctctacact	ttattaatgt	tttcagaaa	tggtcaaaact	28380
aacgctccta	cttttgcttc	tcacgcgggt	gacgcttttt	accatcactt	tttctgccgc	28440
cgccacactc	gaacctcaat	gtttgccacc	ggttgaagtc	tactttgtct	acgtgttgct	28500

gtgctgcgtt agcgtttgcg gtataacatg ttttaccttt gtttttcttc agtgcattga 28560
 ctacttctgg gtcagactct actaccgcag acacgcgcct cagtatcaaa atcaacaaat 28620
 tgccagacta ctcggtctgc catgattgtc ttgtatttta ccttgatttt ttttcacctt 28680
 acttgcgctt gtgattttca cttcactcaa ttttgaaaa cgcaatgctt cgacccgcgc 28740
 ctctccaacg actggatgat ggctcttgca attgccacgc ttggggcggt tggacttttt 28800
 agtgggtttg ctttgcatta caaatttaag actccatgga cacatggctt tctttcagat 28860
 tttccagtta cacctactcc gccgcctccc cgggccatcg acgtgcctca ggttccctca 28920
 ccttctccat ctgtctgcag ctactttcat ctgtaatggc cgacctagaa tttgacggag 28980
 tgcaatctga gcaaagggtc atacacttcc aacgccagtc ggaccgcgaa cgcaaaaaca 29040
 gagagctgca aaccatacaa aacaccacc aatgtaaagc cgggatattt tgtattgtaa 29100
 aacaagctaa gctccactac gagcttctat ctggcaacga ccacgagctc caatacgtgg 29160
 tcgatcagca gcgtaaaacc tgtgtattct taattggagt tccccatt aaagtactc 29220
 aaaccaaggg tgaaaccaag ggaaccataa ggtgctcatg tcacctgtca gaatgccttt 29280
 aactctagt taaaacccta tgtggcttac atgattctat cccctttaat taaataaact 29340
 tactttaaat ctgcaatcac ttcttcgtcc ttgttttgt cgccatccag cagcaccacc 29400
 tccccctctt cccaacttcc atagcatatt ttcgaaaag aggcgtactt tcgccacacc 29460
 ttaaggga cgtttacttc gctttcaagc tctccacga ttttcattgc agatatgaaa 29520
 cgcgccaaag tggaagaagg atttaacccc gtttatccct atggatatcc tactccgact 29580
 gacgtggctc ctccctttgt agcctctgac ggtcttcaag aaaaccacc tggggtcttg 29640
 tccctaaaaa tatccaaacc ttttaactttt aatgcctcca aggcctctaag cctggctatt 29700
 ggtccaggat taaaaattca agatggtaaa ctagtggggg agggacaagc aattcttgca 29760
 aacctgccgc ttcaaatcac caacaacaca atttcactac gttttgggaa cacacttgcc 29820
 ttgaatgaca ataatgaact ccaaaccaca ctaaaatctt catcgccctt taaaatcaca 29880
 gaccagactc tgtcccttaa cataggggac agccttgca ttaaagatga caaactagaa 29940
 agcgctcttc aagcgaccct cccactctcc attagcaaca acaccatcag cctcaacgtg 30000
 ggcaccggac tcaccataaa tggaaacggt ttacaagctg ttcccttaa tgctctaagt 30060
 cccctaacta tttccaacaa taacatcagc ctgcgctatg gcagttccct gacggtgctt 30120
 aacaatgaac tgcaaagcaa cctcacagtt cactccccct taaaactcaa ctccaacaac 30180

tcaattttctc	tcaacactct	atctccgttt	agaatcgaga	atggtttcct	cacgctctat	30240
ttgggaacaa	aatctggctt	gctagttcaa	aacagtggct	taaaagttca	agcgggctac	30300
ggcctgcaag	taacagacac	caatgctctc	acattaagat	atctcgctcc	actgaccatt	30360
ccagactcgg	gctcagaaca	aggcattctt	aaagttaaaca	ctggacaggg	cctaagtgtg	30420
aaccaagctg	gagcgcttga	aacatcccta	ggaggtggat	taaaatatgc	tgataacaaa	30480
ataacctttg	atacaggaaa	cggactgaca	ttatctgaaa	ataaacttgc	agtagctgca	30540
ggtagtggtc	taacttttag	agatgggtgc	ttggtagcca	cgggaaccgc	atttacgcaa	30600
acactgtgga	ctacggctga	tccgtctccc	aactgcacaa	ttatacagga	ccgcgacaca	30660
aaatttactt	tggcgcttac	catttagtgg	agccaagtgc	tggggacggg	ttccattatt	30720
ggagtaaaag	gccccctttc	aagtagcata	ccgtcagcta	ccgttacagt	acaacttaac	30780
tttgattcca	acggagccct	attgagctcc	tcttcaacta	aaggttactg	ggggtatcgc	30840
caaggctccct	caattgaccc	ttaccccata	attaatgcct	taaactttat	gccaaactca	30900
ctggcttata	ccccgggaca	agaaatccaa	gcaaaatgta	acatgtacgt	ttctactttt	30960
ttacgaggaa	atccacaaag	accaatagtt	ttaaacatca	cttttaataa	tcaaaccagc	31020
gggtttttcca	ttagatttac	atggacaaat	ttaaccacag	gagaagcatt	tgcaatgccc	31080
ccatgcactt	tttccctacat	tgctgaacaa	caataaacta	tgtaaccctc	accgttaacc	31140
cgcctccgcc	cttccatttt	attttataaa	ccacccgatc	caccttttca	gcagtaaaca	31200
attgcatgtc	agtaggggca	gtaaaacttt	tgggagttaa	aatccacaca	ggttcttcac	31260
aagctaagcg	aaaatcagtt	acatttataa	aaccatcgct	aacatcggac	aaagacaagc	31320
atgagtccaa	agcttccggg	tctggatcag	atttttgttc	attaacagcg	ggagaaacag	31380
cttctggagg	attttccatc	tccatctcct	tcatcagttc	caccatgtcc	accgtgggtca	31440
tctgggacga	gaacgacagt	tgtcatacac	ctcataagtc	accggctgat	gacgaacgta	31500
cagatctcga	agaatgtcct	gtcgccgcct	ttcggcagca	ctggggccgaa	ggcgaaagcg	31560
cccatgttta	acaatggcca	gcaccgcccc	cttcatcagg	cgcctagttc	ttttagcgca	31620
acagcgcatg	cgcagctcgc	taagactggc	gcaagaaaca	cagcacagaa	ccaccagatt	31680
gttcatgata	ccataagcgt	gctgacacca	gccataacta	acaaattggt	tcactattct	31740
agcatgaatg	tcatactctga	tgttcaagta	aattaaatgg	cgccccctta	tgtaaacact	31800
tcccacgtac	aacacctcct	ttggcatctg	ataattaacc	acctcccgat	accaaataca	31860
tctctgatta	atagtcgccc	cgtacactac	ccgattaaac	caagttgcca	acataatccc	31920

ccctgccata	cactgcaaag	aacctggacg	gctacaatga	cagtgcmaaag	tccacacctc	31980
gttgccatgg	ataactgagg	aaagccttaa	gtcaatagtg	gcacaactaa	tacaaacatg	32040
taaatagtgt	ttcaacaagt	gccactcgta	tgagggtgagt	atcatgtccc	aggggaacggg	32100
ccactccata	aacactgcaa	aaaccaacaca	tcctaccatc	ccccgcacgg	cactcacatc	32160
gtgcatgggtg	ttcatatcac	agtccggaag	ctgaggacaa	ggaaaagtct	cgggagcatt	32220
ttcatagggc	ggtagtgggt	actccttgta	ggggttcagt	cggcaccggg	atctcctcac	32280
cttctggggc	ataacacaca	agttgagatc	tgattttcaag	gtactttctg	aatgaaaacc	32340
aagtgccttc	ccaacaatgt	atccgatgtc	ttcgggtcccc	gcgtcggtag	cgctccttgc	32400
agtacacacg	gaacaaccac	tcacgcaggc	ccagaagaca	gttttccgcg	gacggtgaca	32460
agttaatecc	cctcagtcctc	agagccaata	tagtttcttc	cacagtagca	taggccaaac	32520
ccaaccagga	aacacaagct	ggcäcgtccc	gttcaacggg	aggacaagga	agcagaggca	32580
gaggcatagg	caaagcaaca	gaatttttat	tccaactggt	cacgtagcac	ttcaaacacc	32640
aggtcacgta	aatggcagcg	atcttggggt	tcctgatgga	acataacagc	aagatcaaac	32700
atgagacgat	tctcaagggtg	attaaccaca	gctggaatta	aatcctccac	gcgcacattt	32760
agaaacacca	gcaatacaaa	agcccgggtt	tctccgggat	ctatcatagc	agcacagtca	32820
tcaattagtc	ccaagtaatt	ttcccgtttc	caatctgtta	taatttgcag	aataatgccc	32880
tgtaaatacca	agccggccat	ggcgaaaagc	tcagataatg	cactttccac	gtgcattcgt	32940
aaacacaccc	tcattctgtc	aatccäaaaa	gtcttcttct	tgagaaacct	gtagtaaatt	33000
aagaatcgcc	aggttaggct	cgatgcctac	atcccggagc	ttcattctca	gcattgcactg	33060
caaatgatcc	agcagatcag	aacagcaatt	agcagccagc	tcatccccgg	tttccagttc	33120
cggagttccc	acggcaatta	tcactcgaaa	cgtgggacaa	atcgaaataa	catgagctcc	33180
cacgtgagca	aaagccgtag	ggccagtgca	ataatcacag	aaccagcgga	aaaaagattg	33240
cagctcatgt	ttcaaaaagc	tctgcagatc	aaaattcagc	tcattgcaaat	aacacagtaa	33300
agtttgcggt	atagtaaccg	aaaaccacac	gggtcgacgt	tcaaacatct	cggcttacct	33360
aaaaaagaag	cacattttta	aaccacagtc	gcttcctgaa	caggaggaaa	tatgggtcgg	33420
cgtaaaacca	gacgcgccac	cggatctccg	gcagagccct	gataatacag	ccagctgtgg	33480
ttaaacagca	aaacctttaa	ttcggcaacg	gttgaggctc	ccacataatc	agcggccaca	33540
aaaatcccat	ctcgaacttg	ctcgcgtagg	gagctaaaat	ggccagtata	gccccatggc	33600

acccgaacgc taatctgcaa gtatatgaga gccaccccat tcggcgggat cacaaaatca 33660
 gtcggagaaaa acaacgtata caccgccggac tgcaaaagct gttcaggcaa acgcccctgc 33720
 ggtccctctc ggtacaccag caaagcctcg ggtaaagcag ccattgccaag cgtttaccgt 33780
 gccaaagagcg actcagacga aaaagtgtac tgaggcgctc agagcagcgg ctatatactc 33840
 tacctgtgac gtcaagaacc gaaagtcaaa agttcacccg gcgcgcccga aaaaaccgcg 33900
 gaaaatccac ccaaaaagcc cgcgaaaaac acttccgtat aaaatttccg ggttaccggc 33960
 gcgtcacccg cgcgcgacac gcccgccccg ccccgcgctc ctccccgaaa cccgcgcgcg 34020
 ccacttccgc gttcccaaga caaaggctcg gtaactccgc ccacctcatt tgcattgtaa 34080
 ctcggtcgcc atcttgcggt gttatattga tgatg 34115

<210> 7
 <211> 44
 <212> DNA
 <213> Artificial

<220>
 <223> primer P5L

<400> 7
 gcgcacgcgt ctctatcgat gaattccatt ggtgatggac atgc

44

<210> 8
 <211> 36
 <212> DNA
 <213> Artificial

<220>
 <223> primer P5ITR

<400> 8
 gcgcatttaa atcatcatca ataataacc tcaaac

36

<210> 9
 <211> 31
 <212> DNA
 <213> Artificial

<220>
 <223> primer P5XTOP

<400> 9
 gatacctagg aacgaggagg atttgatatt g

31

<210> 10
 <211> 20

<212> DNA
 <213> Artificial
 <220>
 <223> primer PSXBOT
 <400> 10
 atgtacgect ccgcgctcac

20

<210> 11
 <211> 31
 <212> DNA
 <213> Artificial

<220>
 <223> primer P5E4

<400> 11
 gatcgaattc ccactctgta ccccatctct g

31

<210> 12
 <211> 31967
 <212> DNA
 <213> Simian adenovirus

<220>
 <221> CDS
 <222> (13796) .. (15322)

<220>
 <221> CDS
 <222> (18257) .. (21010)

<220>
 <221> CDS
 <222> (27192) .. (29015)

<400> 12
 catcatcata atatacetta tttgggaacg gtgccaatat gataatgagg aggcgggggtt 60
 aggggtggag tgagggtggg gtgcggatga cgcgggacgc gggcgggggtg ggagtctgac 120
 gtggggcgcg ggggtggagcg cgagggtgag ggcggggcga gggcggcggg cgcggcgga 180
 ttgacgtaca cggtagtaag tttgagcgga aattaagtga attgggcgtg ttttttgtaa 240
 ctttttgacg tacacggtag taagtttgag cggaaattaa gtgaattggg cgtgtttttt 300
 gtaacttttt gacgtacacg gtagtaagtt tgagcggaaa ttaagtgaat tgggcgtggt 360
 ttttgtaact ttttgacgta cacggtagta agtttgagcg gaaattaagt gaattgggcg 420
 tgttttttgt aactttttga cgtacacggg agtaagtttg agcggaaatt aagtgaattg 480

ggcgtgtttt ttgtaacttt ttgacgtaca cggtagtaag tttgagcggg aattaagtga	540
attgggcggtg ttttttgtaa ctttttgacg tacacggtag taagtttgag cggaaattaa	600
gtgaattggg cgtgtttttt gtaacttttt ggtcattttg gcgcgaaaac tgagtaatga	660
ggaagtgaga cggactctgc cctttttttac ggttgggagg gaaaactgct gatcagcgct	720
gaactttggg ctctgacgag gtggtttccc tacgtggcag tgccacgaga aggctcaaag	780
tcctcgtttt attgtgtgct cagccttttt gagggatttt aaacaccgtc agaccgtcaa	840
gaggccactc ttgagtgcga gcgagtagag ttttctctc cgtcgctgcc gcggctgctc	900
agtcttaccg ccaggatgag aatgctgccg gagatcttca ccgggtcctg ggaagatggt	960
ttccagggac ttttagaatc tgaagacaac tttcccaac ctctgagcc ggaggagcta	1020
cctgaggttt cgcttcacga tctgtttgac gtggagggtg agagccccga cggagatccg	1080
aacgaggaag ctgttgatgg tatgttcccc gactggatga tatctcagag cgagagtgct	1140
gaaggcagtg cggactcggg cgtttctggg gttggaaacc tgggtggagg gtatctggac	1200
ttgaagtgtt acgaggaagg ttttctctc agcgactcag agactgatga agcctcagaa	1260
gcggaaggtc aagaggagtc tgtgtgtggt tatgtgaaga ttaatgaggg ggagaacctg	1320
ctggtgttgg actgtccgga ccaacctgga catggctgtc gagcctgtga ctttcaccgg	1380
gggaccagcg gaaaccgga agctatctgt gctttgtgct acatgcgtct gaacgagcac	1440
tgcatataca gtgagtgtta ttcattgggt atttatgggg aaagtgggg gaaagtcttg	1500
agaaggggaa aagtttaaca tgtcattttt gtacttgata ggtccagttt cggacgctga	1560
gggggattct gagtccctg ctggctcttc ccagccctca ccctgctctt tgaccgccac	1620
gcccgcacct gacctagtta gaccaacgcc ctgccgagtg tcctgtagac gacgtgcagc	1680
tgttaattgc atagaagatt tattggcccc tgatgacgag aacgcacctt tgaacctgtg	1740
cctgaaacgc cctaagacat cttgagtgtt tatgctgtta ataaaagtgt tgacccttag	1800
atcctgtgtt tattccttgg gcgtgtgcgc gggatatataa agcagctgag ggctggagtg	1860
ttagtttatt ctgatggagt actggagtga gctgcagaat taccagagcc tccggcgctt	1920
gctggagttg gcctctgcca gaacatccac ctgctggagg ttctgttttg gctcgactct	1980
cagtaacgtg gtgtatcggg tgaagcaaga gtacagctcg cgcttttctg agctgttggc	2040
ccgctacccg gctgtttttg tttctctgga tctaggccat cacgtttatt tccaagaagc	2100
tgtagtcaga tatttggatt tttctactcc cgggcgtgag gtttctgcga ttgccttcac	2160
ctgctttgtg ctagatcgat ggagcgccca aaccgcctg agccgggggt acaccctgga	2220

ctacctgacc atgtccctgt ggagggccat gctgcggaag aggagggctct caggcttctc	2280
gccggcgcg cctccgcacg gactggatcc ggtgctggag gagtcggagc tggaggagga	2340
ggagaacccg agggccggcc tggaccctcc ggcggaatag tgacggaacc ggaggatccc	2400
caagagggta ctagtcaggg gggagggggg ccgaagagaa agcgggatga agaggaggcg	2460
atggaccccg acagggtttct aaaagaactg actttaagct taatgtctaa gagaagacct	2520
gagacgggtgt ggtggctga tttggagaag gagttccacc agggggagat gaatctgttg	2580
tacaagtatg ggtttgagca ggtgaagact cactggctgg aagcctggga ggactgggag	2640
atggctttta acatgtttgc caaggtggcg ctgcgcccg acactattta caccgtgact	2700
aagacgggtgg aaatccgcaa gcctgtgtat gtgattggca acggggccgt ggttcggttc	2760
cagaccaccg accgggtggc ctttaattgc tgtatgcaga acctgggccc ggggggtgatt	2820
aatcttaatg gagtgacctt ttgcaatgtc agattcgcg gggatggatt caacgggacg	2880
gtgtttgccg ccaccacca gataacccta cacggggtgt tcttccagca tgtaggcggg	2940
gcttgtgtag atacctgggc gagggcctct gtgaggggct gcacctttgt gggctgttg	3000
aaagcgggtg tgggtcgacc caagagtgt ctgtctgtga agaaatgtgt gtttgagaga	3060
tgtctgatgg ccatgggtgg ggagggccag ggtaggatcc gccataacgc gggctccgag	3120
aatacctgtt ttgccctgct gaagggtacg gcgaccgtga agcataacat gatctgcggg	3180
gtgggtcact cgcagctgct gacctgtgc gatggcaact gccaggccct gcgcacgggt	3240
catgtggtgt cccaccggcg ccgccctgg ccggtgtttg aacataacat gctgatgcgc	3300
tgtaccatgc acctgggcta ccgcccggc gtgtttgtgc cccatcagt taacctgacc	3360
cacaccaagg tgttgctgga gacggatgct ttttcgcgag tgaatctgaa tgggggtgtc	3420
gatctgacta tggagatgta caagatagt agatttgatg aatcaaagac ccgttgtcgc	3480
ccctgcgagt gcggtgcaa tcacctgagg atgtatccc tgacctgaa cgtgacggag	3540
gagctgcgcc cggaccacca gatgctgtcc tgtctgcgca ccgattacga aagcagcgat	3600
gaggattaag aggtgagggg cggggcttgc atgggggtata aagggtggggg aggaggtggg	3660
gagggggaaa acccaaatg agcggatcga tggaggggag cgctgtgagt tttgagggcg	3720
gggtgttcag cccatatctg acaaccgct tccccgctg ggcaggagt cgtcagaatg	3780
tgggtgggtc caacgtggac ggacgtccgg tggccctgc caactccgcc actctcacct	3840
acgccaccgt cggatcgtcg ctggacaccg ccgctgccgc cgcgcttca gccgccgtt	3900

ctactgctcg	cggtatggca	gctgatttcg	gactgtatca	gcaactggct	gcgcctcgct	3960
cgtcgctgag	agaagatgat	gccctgtccg	tggtgctgac	ccgcctggag	gagctgtccc	4020
agcagctgca	agagctgtct	gccaaagtgg	atgcacagaa	cgcccccgct	acccaatgaa	4080
taaataaacg	agacaccgag	tgtgttttga	aatcaaaatg	tgtttttatt	tgttttttct	4140
ggcgcggtag	gcccttgacc	acctgtcgcg	gtcgttaagg	accttgaggga	tgtttttccag	4200
cacccggtag	aggtgggctt	ggatggtgag	gtacatgggc	atgagcccg	ctcggggggtg	4260
gaggtagcac	cactggaggg	cgtcgtgctc	gggggtgggtg	ttgtagataa	tccagtcgta	4320
gcagggtttt	tgggcatgga	agcggaagat	gtctttgaga	agcaggctga	tggccagggg	4380
gaggcccttg	gtgtaggtgt	tcacaaagcg	gttgagctgg	gagggatgca	tgcggggggga	4440
gatgagatgc	atcttggcct	gaatcttgag	gttggcgatg	ttgccgccc	gatcccgccg	4500
ggggctcatg	ttgtgcagga	ccaccaggac	ggtgtagccg	gtgcacttgg	ggaatttgct	4560
atgcaacttg	gaaggggaagg	cgagggaagaa	cttgagagacc	cccttggtgc	cgccgaggtt	4620
ctccatgcat	tcgtccatga	tgatggcgat	gggacccctg	gcggccgccc	tggcgaagac	4680
gttgtcgggg	tgggagacgt	cgtagttctg	ttccagggtg	agctcgtcgt	aggccatttt	4740
gacgaagcgg	gggagcaggg	tgctcgactg	ggggacgatg	gtaccttcgg	gacccggggc	4800
gtagttgcc	tcgcagattt	gcatctccca	ggccttgatc	tccgaggggg	ggatcatgtc	4860
cacctggggc	gcgatgaaga	agacggtctc	cggggcgggg	ttgatgagct	gggaggagag	4920
gaggttgagg	agcagctgcg	acttgccgca	cccggtgggc	ccgtagatga	ccccgatgac	4980
gggttgagc	tggtagttta	aggagctgca	gctgccgtcc	tcgcgcagga	acggggcgac	5040
ctcgttcac	atgcttctga	cgtagtggtt	ttccctgacg	aggctcttga	agagccgctc	5100
gccgcccagg	gagagaagct	cttccaggct	gcggaaatgc	ttgaggggtt	tgaggccgct	5160
ggccatggtc	atcttttcca	gggactggcg	gagcaggtac	aggcggtccc	agagctcggg	5220
gacgtgttct	acggcatctc	gatccagcag	acttcttggt	tgcgggggtt	ggggcggtt	5280
tggctgtagg	ggaccagccg	gtgcgcgtcc	agggaggcga	gggtgacgtc	tttccagggc	5340
cgcagcgttc	gcgtgagggg	ggtctcgggtg	acgggtgaagg	gatgcgctcc	cggttgggcg	5400
ctggccaggg	tcctcttgag	actcatcctg	ctggtgtgga	agcgggctgc	ttctccctgg	5460
gagtcggcca	ggtagcattt	gagcatgagg	tcgtagctga	gggcctcggc	cgcgtggccc	5520
ttggcgcgca	gcttgccctt	ggagacgtgt	ccgcaggcgg	gacagtgcag	gcacttgagg	5580
gcgtagagct	tgggggcccag	gaagacggac	tcgggggagt	aggcgctggc	gccgcactga	5640

gcgcacgtgg tctcgactc gacgagccag gtgagctccg ggtgttgggg atcaaaaacc	5700
agctggcccc cgtgtttttt gatgcgcttc ttacctcggg tctccatgag gcggcgctccg	5760
gcttcggtga cgaagaggct gtccgtgtcg ccgtagacgg atttgagcgc gcgctgctcc	5820
aggggaatcc cgcgacctc cgcgtgcagg aactcggacc actctgagac gaaggcccgg	5880
gtccacgcga ggacaaagga ggcgatctgg gacgggtagc ggtcgttctc caccagggga	5940
tccaccttct ccagggtgtg caggcagagg tcgtcctcct ccgcgtccat gaaggtgatt	6000
ggcttgtaag tgtatgtcac gtgaccgtcg ggtcgcgcg tgggcttata aaagggggcg	6060
tgcccggcct ccccgtcact ttcttcgca tcgctgtgga cgagatccag ctgctcgggt	6120
gagtaggcgc gctggaaggc gggcatgacc tcggcgctga ggtgtcagt ttccacgaac	6180
gaggtggatt tgatattgac ctgtccggcg gcgatgcttt tgacgggtggc ggggtccatc	6240
tggtcagaaa agacgatctt ttgttgtcc agcttgggtg cgaacgacct gtagagggcg	6300
ttggagagca gcttggcgat ggagcgagg gtctggttct tctcgcggtc ggcgcgctcc	6360
ttggcggcga tgttgagctg gacgtactcg cgggccacgc agcgccattc ggggaagacg	6420
gtggcgcgct cgtccggcag gaggcgacg cgcagccgc ggttgtgcag ggtgatgagg	6480
tccacgctgg tggccacctc gccgcgcagt ggctcggtgg tccagcagag gcgcccggcc	6540
ttgcgcgagc agaagggggg caggacgtcg agctggctct ccgcgggggg gtccggcgctc	6600
atggtgaaga tggccggtag cagggtggcg tcgaagtagt cgatggcgac cgcggggctc	6660
gcgagggcgc gttcccagtc cctgaccgcc agggcgcgct cgtaggggtt gaggggcgcc	6720
ccccagggca tgggatgggt gagggccgag gcgtacatgc cgcagatgtc gtagacgtag	6780
aggggctcgc ggagcacgcc gaggtagggt ggatagcagc gtccgccgcg gatgctggcg	6840
cgcacgtagt cgtacatctc gtgcgagggg gcgaggaggc cgcctccgag gtcgccgcgc	6900
tgcggtctga cggcccggta ggtgacctgg cggaagatgg cgtgcgagtt ggaggagatg	6960
gtgggccgct ggaagatgtt gaagctggcc tcggggagtc cgacggcgtc gtggacgaac	7020
tgggcgtagg agtcgcgcag cttctgcacg agcgggcggt tgacgagcac gtccagggcg	7080
cagtagtcga ggtctcgcg gacgaggtcg taacggggct cttgcttctt ttcccagagt	7140
tcgcggttga ggaggtactc ctgcgcatcc ttccagtact cttcgggcgg aaagccgcgt	7200
tcgtccgcca ggtagaacc cagcatgtag aagcggttga cggctcggtg gggacagcag	7260
cccttctcga cgggcaggga gtaggcctgc gcggccttcc tgagcgaggt gtgggtgagg	7320

gcgaaggtgt	cgcgccaccat	gaccttgagg	aactggaacc	tgaagtcggt	gtcgtcgcag	7380
gcgccccgct	cccagagccc	gtagtcggtg	cgtttctggc	tgcgsggggtt	gggcagggcg	7440
aaggtgacgt	cgttgaagag	gatcttgccg	gcgcgcggca	tgaagttgcg	ggtgatcctg	7500
aagggccccc	gcacgtccga	gcggttggtta	atgacctggg	ccgcgaggac	gatctcgtcg	7560
aagccgttga	tggtgtggcc	gacgatgtag	agctcgacga	agcgsgggcg	cccctgcagc	7620
ttggggggcct	tcttgagctc	ctcgtagggtg	aggcagtcgg	gcgagtagag	gcccagctcc	7680
tgtcgggccc	attcgggccac	ctggggggtt	gcttgcaaga	agccccgcca	gagctgcagg	7740
gcgagctggg	tctggaggcg	gtcgcggtag	tcgcggaact	ttttgcccac	cgccatcttc	7800
tcgggggtga	ccacgtagaa	ggtgcggccg	tcctggcccc	aggcgcccca	gttctgctcg	7860
cgggcgagac	ggcaggccctc	ctcgacgagg	gcctcctccc	cggagagatg	catgactagc	7920
atgaagggga	cgagttgctt	gccgaaggca	cccattccacg	tgtaggtctc	tacgtcgtag	7980
gtgacgaaga	gacgttcggt	gcgaggatgc	gagccgagag	gaaagaagtt	gatctcctgc	8040
caccagcccg	aggagtgggc	gttgacgtgg	tggaagtaga	agtcacgccg	gcggaccgtg	8100
cattcgtgct	gatatttgta	aaagcggggc	cagtactcgc	agcgctgcac	gctctgcact	8160
tcctgaacga	gatgcacccg	gcgcccgcgc	accaggaggc	ggagggggca	gtccagtgga	8220
gcttcgggcgc	gctgtccttc	agcctcgtca	tgtcttctct	cacctgcacg	ctcctgctgt	8280
gggtggagga	cggagggagt	gacgacgccg	cgcgagccgc	aggtccagat	gtcgacgcgc	8340
ggcggcctga	ggctcagcgc	caggggtgcg	atctgagcgg	cgtccagggg	gtcgaggaag	8400
gcctcgctga	ggtcgacggg	cagcgtccgc	cggtggactt	gcaggagacg	ggtaagggcc	8460
ggcgccaggc	gctgatggta	cttgatctcg	agcggttcgt	tggtggagggt	gtcgatggcg	8520
tagagcaggg	cctgaccgcg	ggcggcgacg	atgggtgccg	ggtgccggcg	gtaggtggcg	8580
tattcggggg	ggctcgttac	atcacccgcc	tgggcctggc	gccgggcggc	agcgggggtt	8640
ctggtcccgc	cggcatgggc	ggcagcggca	cgtcggcgcg	gggctccggc	agcggctgggt	8700
gctgagctcg	cagctgactg	gcgtgcgcga	cgacgcggcg	ggtgaggtcc	tggatgtgcc	8760
tcgctgctg	gaagaccacc	gggtccccga	ctcggaacct	gaaagagagt	tcgacagaat	8820
caatctcggc	atcgttgacg	gccgcctgac	gcaggatctc	ctgcacgtcg	cccaggttgt	8880
cctggtaggc	gatctcggac	atgaactggt	cgatctcttc	ctcctggagt	tcgccgcgtc	8940
cggcgcgctt	gacgggtggc	gcgaggtcgt	tggagatgcg	agccatgagc	tgggagaagg	9000
cgttgaggcc	gttctcgttc	cacacgcgac	tgtagacgac	ggtgccgacg	gcgtccccgg	9060

cgcgcatgac cacctgcgcg acgttgagct ccacgtgtcg cgcgaagacg gcgtagtgtc	9120
gcaggcgctg gaagaggtag ttgaggggtg tggcgatgtg ctgcagacg aagaagtaca	9180
tgaccacagc ggcagcgtc atctcgttga tgtctccgag ggcttccaag cgctccatgg	9240
cctcgtagaa gtcgacggcg aagttgaaga actgggagtt gcgcgccgcg accgtcagct	9300
cgtcttgcga gagccggatc agctgggcca cggctctccc cacctcgcgt tcgaaggccc	9360
cggcgcttc ttctctctct gggtctctcg cggcctcttc ttccatgacg gcttctctct	9420
cctccgggtc ctccggcagc ggcctccggc ggcgacggcg cctgatgggc aggcgggtcca	9480
cgaagcggtc gatgatctct ccgcggcggc ggcgcatggt ttccggtgacg gcgcggccgt	9540
tctctcgggg ccgcagttcg aagacgcccc cgcgcaggcc gccggcgccg ccgagagggg	9600
gcaggaggtg ggggccttcg ggcagcgaga gggcgctgac gatgcaccgt atcatctgtt	9660
gcgtaggtac agctctccag gagtctgtga gcgagtcag ttggacggga tccgagaact	9720
tttcgaggaa agcttcgatc caatcgcagt cgcaaggtaa gctgaggacg gtgggatgag	9780
gggcttggcg ggaggcgagc gcggcagaag aggaggagga gggcaggctg gaggtgatgc	9840
tgctgatgat gtaattgaag taggcggttt tcaaaccggc gatgggtggcg aggaggacga	9900
cgtctttggg cccggcctgc tggatgcga ggcggtcggc catgccccag gcgtggctct	9960
ggcatcggcg caggtccttg tagtagtctt gcatgagtct ctgcacgggg acgtcgtctt	10020
cgtcggcccc gtcggccatg cgggtggagc cgaaccgcg caggggctgc agcagggccca	10080
ggtcggcgac caccggttcg gccagcacgg cctgctggat ctgggtgagc gtgggtctgga	10140
agtcgtccag gtccacgaag cggtggtagg agcccgtgtt gatggtgtag gtgcagttgg	10200
ccatgacgga ccagttgacg acttgcatgc cgggtctggg gatctcggtg tagcggaggc	10260
gcgagtaggc ccgcgactcg aagacgtagt cgttgacagg gcgcacgagg tactggtagc	10320
cgacgaggaa gtgcggcggc ggctcgcggc agaggggcca gcgcacggtg gcgggggcgc	10380
cggggggccag gtcctccagc atgaggcggc ggtagtggta gacgtagcgc gagagccagg	10440
tgatgccggc ggcggaggtg gcggcgcggg cgaagtcgcg gacgcgggtc cagatgttgc	10500
gcaagggggc gaagcgtcc atggtgggca cgctctggcc ggtgaggcgg gcgcagtcct	10560
gcacgctcta gacgggacag agagcgggag gttagcggct ccgctccgtg gcctggggga	10620
cagaccgcca ggggtcgacg gcggggaacc ccggttcgag accggtgga tccgtccgtc	10680
cccgcgcgc cggccccgcg tccacgacct caccagaggc cgagaccag ccgcgggtgc	10740

cggaccccag atacggaggg gagccttttt gtggtttttt cccgtagatg catccggtgt 10800
 tgcgacagat gcgtccgtcg ccagcgcgcg cgacgcagcc gccgctcccg cccccacta 10860
 gcgccgcgga ggctctgtcc ggcggccgcg gcgacccgga ggaggaggcc atcctcgact 10920
 tggagaagg cgagggcctg gcccggtcg gagcgccctc ccccgagcgc catccccgcg 10980
 tgcagctggc gagagactcg cgccaggcct acgtgccgcc gcagaatctg ttgagggacc 11040
 gcagcggcca ggagcccag gagatgaggg accgcagggt tcacgcgggg cgggagctgc 11100
 gcgcgggctt cgaccgtcgg cgggtgttgc gcgccgaaga cttcgagccc gacgagcgca 11160
 gcggagtaag tccggcacgg gcgcacgtgt cggcgcccaa cctggtgacc gcgtacgagc 11220
 agacggtgaa cgaggagcgg agctttcaga aaagcttcaa caaccacgtg cgcaccctga 11280
 tcgcgcgca ggaggtggcc atcggcctga tgcattctgt ggactttgtg gaggcgtagc 11340
 tgcagaacct gtcgagcaag ccgctgacgg cgcagttgtt cctgatcgtg cagcacagtc 11400
 gggacaacga gacgttccgc gaggcgatgc tgaacatcgc ggagcccag ggcgctggc 11460
 tcttggaact gattaacatc ctgcagagca tcgtggtgca ggagcgcagc ctgagcctgg 11520
 ccgacaaggt ggcggccatc aactacagca tgttgagcct gggcaagttt tacgcccga 11580
 agatctacaa gagcccctac gtgcccatac acaaggaggt gaagatcgac agcttttaca 11640
 tgcggatggc gctgaaagtg ctgacgctga gcgacgatct ggggggtgtac cgcaacgacc 11700
 gcatccacaa ggccgtgagc gccagccgcc ggcgcgagct gagcgaccgc gagctgatgc 11760
 acagcctgcg gagggcgctg gcgggcgcgg gcggcgccga ggaggccgag tccctacttcg 11820
 acatgggggc ggacttgag tggcagccca gcgcgcgggc cctggaggcg gcgggctacc 11880
 gcggcgccgg cggcgtggtc gaggcggagg acgaggacga ggtggagtac gaggaggagg 11940
 actgatcggc gaggtgtttt cgtagatgca gcgcgcgacg gcggcgccga gcgggcccga 12000
 gggggacccc gccgtgctgg cggccctgca gagccaacct tcgggcgtga acgcctccga 12060
 tgactgggcy gcggccatgg accgcatttt ggccttgacc acccgcaacc ccgaggcctt 12120
 tagacagcag ccgcaggcca accgcttttc ggccatcttg gaagccgtgg tgccctcgcg 12180
 caccaacccc acgcacgaga aggtccctggc ggtggtgaac gcgctgctgg agagcaaggc 12240
 gatccgcaag gacgaggcgg ggctgattta caacgccctg ctggagcggg tggcgcgcta 12300
 caacagcacc aacgtgcagg ccaacctgga ccgtctgacg acggacgtgc gggaggcggt 12360
 ggcgcagcgg gagcgcttca tgcgcgacac gaacctgggc tcgcaggtgg ccctgaacgc 12420
 cttcctgagc acgcagccgg ccaacgtgcc gcgcgggcag gaggactacg tcagtttcat 12480

cagcgcgctg cgcctcctgg tggccgaggt gccgcagagc gaggtgtacc agtcgggtcc 12540
ggactacttc ttccagacct cgcggcaggg cctgcagacg gtgaacctga cgcaggcctt 12600
caagaacctg gaaggcatgt ggggcgtgcg gggccccgtg ggcgaccggg cgacgatctc 12660
cagcttgctg acgccgaaca cgcggctgct gctgctgctg atcgcgccct tcaccaatag 12720
cagtaccatc agccgcgact cgtacctggg ccacctgac acgctgtacc gcgaggccat 12780
cgggcaggcg caggtggacg agcagacctt ccaggagatt acgagcgtga gccggggccct 12840
ggggcagcag gacacgggta gcctggaggc gacgctgaat tttctgctga ccaaccggcg 12900
gcagaagatc ccctcccagt acacgctgag cacggaggag gagcgcacatc tgcgctacgt 12960
gcagcagtc gtgagcctgt atctgatgcg cgagggggcg agcccctcgt cggcgctgga 13020
catgacggcc cgtaacatgg agccgtcgct gtacgcggcc caccggccgt tcgtgaaccg 13080
cctgatggac tacctgcacc gcgccgccgc catgaacggc gactacttta cgaacgccat 13140
cctgaacccg cactggatgc cgcgctcggg tttctacacg ggggactttg acatgccccga 13200
gggcgacgac gggttcctgt gggacgacgt gtcggacagc gtgttcgcgc cggtgcgctc 13260
gggcaagaag gagggcgggc acgagctgcc gctgtccgtg gtggaggcgg cgtcgcgcg 13320
ccagagcccg ttcccagcc tcccgctcgt gtcggcgagc agcagcagcg gccgggtctc 13380
gcgcccgcgg ctggaggcg actacctgaa cgaccgcgtg ctgcgccccg cccggcccaa 13440
gaactttccc aacaacgggg tggagagcct agtggataag atgaatcgct ggaagacct 13500
cgcccaggag cagcgggagt gggaggagag tcagccccgc cccctgcctc cgcgcgcctc 13560
caggtggcgc cggcgggaag aagaccgga agactcggcg gacgatagca gcgtgttgg 13620
cttggggggg accgggtgcc cctcgacaaa cccgttcgcc cacctgcgcc cgcaggggccg 13680
gctgggtcgg ctgtattgag gaaagaaact aataaaagaa aaaagagctt gcttaccaga 13740
gccatggtcg cagcgtcggg ccccttgtgt gtgttttctc ctccccggta gcgaa atg 13798
Met
1
agg cgc gcg gtg gga gtg ccg ccg gtg atg gcg tac gcc gag ggt cct 13846
Arg Arg Ala Val Gly Val Pro Pro Val Met Ala Tyr Ala Glu Gly Pro
5 10 15
cct cct tct tac gaa acg gtg atg ggc gcc gcg gat tcg ccg gcc acg 13894
Pro Pro Ser Tyr Glu Thr Val Met Gly Ala Ala Asp Ser Pro Ala Thr
20 25 30
ctg gag gcg ctc tac gtc cct ccc cgc tac ctg ggg cct acg gag ggg 13942
Leu Glu Ala Leu Tyr Val Pro Pro Arg Tyr Leu Gly Pro Thr Glu Gly

35	40	45	
agg aac agc atc cgt tac tca gag ctg gcg ccg ctg tac gac acc acc Arg Asn Ser Ile Arg Tyr Ser Glu Leu Ala Pro Leu Tyr Asp Thr Thr 50 55 60 65			13990
cgc gtg tac ctg gtg gat aac aag tcg gcg gac atc gcg tcg ctg aac Arg Val Tyr Leu Val Asp Asn Lys Ser Ser Ala Ser Leu Asn 70 75 80			14038
tac cag aac gac cat agc aac ttt ctg acc acg gtg gtg cag aac aat Tyr Gln Asn Asp His Ser Asn Phe Leu Thr Thr Val Val Gln Asn Asn 85 90 95			14086
gac ttt acc ccg gtg gag gcg ggc acg cag acc ata aat ttc gac gag Asp Phe Thr Pro Val Glu Ala Gly Thr Gln Thr Ile Asn Phe Asp Glu 100 105 110			14134
cgc tcg cgg tgg ggc ggc gac ctg aaa acc atc ctg cgc acc aac atg Arg Ser Arg Trp Gly Gly Asp Leu Lys Thr Ile Leu Arg Thr Asn Met 115 120 125			14182
ccc aac atc aac gag ttc atg tcc acc aac aag ttc agg gcc cgg ttg Pro Asn Ile Asn Glu Phe Met Ser Thr Asn Lys Phe Arg Ala Arg Leu 130 135 140 145			14230
atg gta gag aaa gtg aac aag gaa acc aat gcc cct cga tac gag tgg Met Val Glu Lys Val Asn Lys Glu Thr Asn Ala Pro Arg Tyr Glu Trp 150 155 160			14278
ttt gag ttc acc ctg ccc gag ggc aac tac tcg gag acc atg acc ata Phe Glu Phe Thr Leu Pro Glu Gly Asn Tyr Ser Glu Thr Met Thr Ile 165 170 175			14326
gac ctg atg aat aac gcg atc gtg gac aac tac ttg gaa gtg ggg cgg Asp Leu Met Asn Asn Ala Ile Val Asp Asn Tyr Leu Glu Val Gly Arg 180 185 190			14374
cag aac ggg gtg ctg gag agc gac atc ggg gtg aag ttt gac acg cgc Gln Asn Gly Val Leu Glu Ser Asp Ile Gly Val Lys Phe Asp Thr Arg 195 200 205			14422
aac ttc cgg ctg ggc tgg gac ccg gtc acc aag ctg gtc atg ccc ggc Asn Phe Arg Leu Gly Trp Asp Pro Val Thr Lys Leu Val Met Pro Gly 210 215 220 225			14470
gtg tac acc aac gag gcc ttc cac ccc gac atc gtc ctg ctg ccc ggc Val Tyr Thr Asn Glu Ala Phe His Pro Asp Ile Val Leu Leu Pro Gly 230 235 240			14518
tgc ggc gtg gac ttc acg cag agc cgg ctg agc aac ctg ctg ggg atc Cys Gly Val Asp Phe Thr Gln Ser Arg Leu Ser Asn Leu Leu Gly Ile 245 250 255			14566
cgc aag cgg atg ccc ttc cag gcg ggt ttt cag atc atg tac gag gac Arg Lys Arg Met Pro Phe Gln Ala Gly Phe Gln Ile Met Tyr Glu Asp 260 265 270			14614

ctg gag ggc ggc aac atc ccc gcc ttg cta gac gtg gcg aaa tac gag Leu Glu Gly Gly Asn Ile Pro Ala Leu Leu Asp Val Ala Lys Tyr Glu 275 280 285	14662
gcc agc att cag aag gcg cgg gag cag gcc cag gag atc cgc ggc gac Ala Ser Ile Gln Lys Ala Arg Glu Gln Gly Gln Glu Ile Arg Gly Asp 290 295 300 305	14710
aac ttt acc gtc atc ccc cgg gac gtg gag atc gtg ccc gtg gag aag Asn Phe Thr Val Ile Pro Arg Asp Val Glu Ile Val Pro Val Glu Lys 310 315 320	14758
gat agc aag gac cgc agt tac aac cta ctc ccc gcc gac cag acc aac Asp Ser Lys Asp Arg Ser Tyr Asn Leu Leu Pro Gly Asp Gln Thr Asn 325 330 335	14806
acg gcc tac cgc agc tgg ttc ctg gcc tac aac tac gcc gac ccc gag Thr Ala Tyr Arg Ser Trp Phe Leu Ala Tyr Asn Tyr Gly Asp Pro Glu 340 345 350	14854
aag gcc gtc agg tcc tgg acg ctg ctg acc acc acg gac gtc acc tgc Lys Gly Val Arg Ser Trp Thr Leu Leu Thr Thr Thr Asp Val Thr Cys 355 360 365	14902
ggc tcg cag cag gtg tac tgg tcg ctc ccg gac atg atg caa gac ccc Gly Ser Gln Gln Val Tyr Trp Ser Leu Pro Asp Met Met Gln Asp Pro 370 375 380 385	14950
gtg acc ttc cgg ccc tcc agc caa gtc agc aac tac ccc gtg gtg gga Val Thr Phe Arg Pro Ser Ser Gln Val Ser Asn Tyr Pro Val Val Gly 390 395 400	14998
gtc gag ctc ctg ccg gtg cac gcc aag agc ttt tac aac gag cag gcc Val Glu Leu Leu Pro Val His Ala Lys Ser Phe Tyr Asn Glu Gln Ala 405 410 415	15046
gtc tac tcg cag ctc atc cgc cag tcc acc gcg ctc acg cac gtc ttc Val Tyr Ser Gln Leu Ile Arg Gln Ser Thr Ala Leu Thr His Val Phe 420 425 430	15094
aac cgc ttc ccc gag aac cag atc ctg gtg cgc ccg ccc gct ccg acc Asn Arg Phe Pro Glu Asn Gln Ile Leu Val Arg Pro Pro Ala Pro Thr 435 440 445	15142
att acc acc gtc agt gaa aac gtt ccc gcc ctc aca gat cac gga acc Ile Thr Thr Val Ser Glu Asn Val Pro Ala Leu Thr Asp His Gly Thr 450 455 460 465	15190
ctg ccg ctg cgc agc agt atc agt gga gtc cag cgc gtg acc atc act Leu Pro Leu Arg Ser Ser Ile Ser Gly Val Gln Arg Val Thr Ile Thr 470 475 480	15238
gac gcc cgg cga agg acc tgc ccc tac gtg cac aag gcc ctg gcc ata Asp Ala Arg Arg Arg Thr Cys Pro Tyr Val His Lys Ala Leu Gly Ile 485 490 495	15286

gtc gct ccc aaa gtg ctc tct agc cgc acc ttt taa caagcatgtc	15332
Val Ala Pro Lys Val Leu Ser Ser Arg Thr Phe	
500 505	
cattctcatc tcgcccagaca acaacacagg ctggggcctg cgctcggcgg gcatgtacgg	15392
cggcgccaag cggcgctcca gcgagcacc cgtccgcgtc cgcggccact accggggccc	15452
ctggggcgcc cacaagcgcg gcgtctccac gcgcaccacc gtgcagcagc ccacgcagcg	15512
cgtcgtggcc caggccagac gctaccgccc gcccaagtcg acgggtggacg ccgtcatcga	15572
cagcgtggtg gccgacgcgc ggcgatacgc tcgacgcaag cggcgctctgc accgcgctcg	15632
ccgtcccacc gccgccatgc tggcgccag agcggtcctg agacgcgcgc gccgcgtggg	15692
ccgccgagcc atgcgccgag ccgcggccaa cgcagcgcg ggtcgcgccc gtcgtcaggc	15752
cgcggcgag gccgccgccc ccacgcgcaa cctggcccaa cccgcgggg gaaacgtgta	15812
ctgggtgcca gacgcgtcgg gcgtgcgcgt gccggtgcgc accgcggccc ctcggagtta	15872
gaagacaaaa agacggacga agactgagtt tccctgtcgt tgccagcatg agcaagcgca	15932
agttcaaaga agagctgctg gaggccctcg tgcccagat ctacggcccg gccgcgctg	15992
ccgcgcgggt ggcgagcgtc aagcccgaag ttaagcccc cgcgtgaag cgggttaaaa	16052
agcgggaaaa gaaagaggag aagcaggaag cagggttgcg agacgtcgac gacggcggtg	16112
agttcgtgcg gtccttcgcg ccccgctgcc ggggtgcagtg gcggggtcgc cgcgtcaagc	16172
tcgtcccgcg gccgggcacc gtggtgtctt tcacccccgg cctgcgttcg gccacgcgcg	16232
gcctgaagcg cgagtacgac gaggtctatg gcgacgaaga catcctggag caggccgccc	16292
agcagctcgg ggagtttgcg tacggcaagc gcggccgcta cggggaggtg gcgctggcgc	16352
tggaccaggg caatcccacg cccagcctca agcccgtcac gctgcagcag gtgctgcccg	16412
tgagcgcgtc gaccgagagc aagcggggca tcaagagggga gatgggagc ctcagacca	16472
ccatgcaact catggtgccc aaacggcaga agctggagga cgtgctggag aacatgaaag	16532
tggatcccag catcgagccc gaagtgaag tgcgacccat caaggaagtg gggccgggccc	16592
taggcgtgca gacggtggac attcagatcc ccgtgcgcgc ccccccggt tctgccacca	16652
ctacgacggc cgtggaggcc atggaaacgc agacggagct gccgcgggccc ttggcggcag	16712
ccgccaccgc cgcgcgggt acccgagaga tgggcatgca gaccgacccc tggtagcagt	16772
tcgcgggccc cgcgcgtcgt ccacgagccc gtcggtaacg ggcgaccacc tcccggctcc	16832
ctgactacgt cttgcatcct tccatcacgc cgacgcccgg ctaccgcgga acgaccttcc	16892
gccccgggtc cgcgcgcacc accaccgccc gtcgtcgcac caccgcggc cgtcgcagcc	16952

gtcgcgcact ggctcccatc gcggttcgcc gcgtcgctccg ccgggggtcgc acgctgaccc 17012
 tgcccaccgc gcgttaccac cccagcatcg tcatttaacc tgcgctgccg ttttgcagat 17072
 ggctctgacg tgccgctttc gcttccccgt tcggcactac cgaggaagat ctgcgcgtag 17132
 gactgggtcta gcgggcagcg gtctdcgacg ccgccgccgc gcggtgcacc ggcgcagtaa 17192
 gggcggcatt ctgcccgcgc tgatccccat tatcgccgcc gccatcgggg cgatccccgg 17252
 cgtggcctcg gtggccttgc aagcagctcg caaaaattaa ataaagaagg cttgacactc 17312
 actgcctggt cctgactggt tcatgcagac aagacatgga agacatcaat tttgcgtcgt 17372
 tggccccgcg gcacggctcg cggccgttca tgggcacctg gaacgagatc ggcaccagcc 17432
 agctcaacgg gggcgctttc agttggagca gcctgtggag cggcattaaa aactttgggt 17492
 ccacgattaa gacctatggc aacaaggcgt ggaacagtag cactggtcag atgctccgcg 17552
 ataagctgaa ggaccagaac ttccagcaga aagtggtaga cggctctggcc tcgggcatca 17612
 acgggggtggt ggacctggcc aaccaggcgg tgcagaacca gatcaaccag cgtctggaga 17672
 acagccgccca gccgcccgcg gccctgcagc agcgtccgca ggtggaggag gtggaagtgg 17732
 aggagaagct gccgcccctg gagacggtgt cgccggtggg cgtgcctagc aagggggaga 17792
 agcggccgcg gcccgagctc gaggagacct tagtgaccga gacctggag ccgcccctcgt 17852
 acgagcaggc cttgaaagag gggggccacgc ccctgccc atcgccgccc atcgccacca 17912
 tggcccgacc ggtctacggc aaggaaacaca aagccgtgac gctagagctg cctccgcccg 17972
 cgcccaccgt acccccgatg cccggtccca ccctgggcac cgccgtgcct cgtcccgcgg 18032
 ccccgccggt cgccgtggcc acgcccgcgc gcccgagtcg cggagccaac tggcagagca 18092
 ctctgaacag catcgtgggc ctgggagtga aaagcctgaa acgcccgcgg tgttactatt 18152
 aaagccagct aaatacccat gtgttgtatg cgccctcctgt gtcacgccag aaaaagccag 18212
 ccgagtgcg ggtcaccgcc gccgccaaga gcgcgccttt caag atg gcc acc ccc 18268
 Met Ala Thr Pro
 510
 tcg atg atg ccg cag tgg tct tac atg cac atc gcc ggg cag gac gcc 18316
 Ser Met Met Pro Gln Trp Ser Tyr Met His Ile Ala Gly Gln Asp Ala
 515 520 525
 tcg gag tac ctg agc ccg ggc ctg gtg cag ttc gcc cgc gcc acc gac 18364
 Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala Arg Ala Thr Asp
 530 535 540
 acg tac ttc agc ctg ggc aac aag ttt agg aac ccc acg gtg gcc ccc 18412
 Thr Tyr Phe Ser Leu Gly Asn Lys Phe Arg Asn Pro Thr Val Ala Pro

545	550	555	560	
acc cac gac gtg acg acg gac cgg tcc cag cgg ctg acg ctg cgg ttc				18460
Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu Thr Leu Arg Phe	565	570	575	
gtg ccc gtc gac cgc gag gac acc cgc tac tgc tac aaa gtg cgc ttc				18508
Val Pro Val Asp Arg Glu Asp Thr Ala Tyr Ser Tyr Lys Val Arg Phe	580	585	590	
acg ctg gcc gtg ggc gac aac cgc gtg ctg gac atg gcc agc acg tac				18556
Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met Ala Ser Thr Tyr	595	600	605	
ttt gac atc cgc ggc gtg ttg gac cgc ggt ccc agc ttc aaa ccc tac				18604
Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser Phe Lys Pro Tyr	610	615	620	
tcc ggc acc gcc tac aac tcc ctg gcc ccc aag ggc gcc ccc aac cgc				18652
Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly Ala Pro Asn Pro	625	630	635	640
tca gaa tgg aag ggc tca gac aac aaa att agt gta aga ggt cag gct				18700
Ser Glu Trp Lys Gly Ser Asp Asn Lys Ile Ser Val Arg Gly Gln Ala	645	650	655	
ccg ttt ttt agt aca tcc att aca aag gat ggt att caa gtg gcc act				18748
Pro Phe Phe Ser Thr Ser Ile Thr Lys Asp Gly Ile Gln Val Ala Thr	660	665	670	
gat act tct agc gga gct gtg tat gct aaa aag gaa tat cag cct gaa				18796
Asp Thr Ser Ser Gly Ala Val Tyr Ala Lys Lys Glu Tyr Gln Pro Glu	675	680	685	
cca caa gta ggg caa gaa caa tgg aac agc gaa gcc agt gat agt gat				18844
Pro Gln Val Gly Gln Glu Gln Trp Asn Ser Glu Ala Ser Asp Ser Asp	690	695	700	
aaa gta gct ggt agg att cta aaa gac aca aca ccc atg ttc cct tgt				18892
Lys Val Ala Gly Arg Ile Leu Lys Asp Thr Thr Pro Met Phe Pro Cys	705	710	715	720
tac ggt tcc tac gcc aag ccc aca aat gaa cag ggg ggg caa ggc act				18940
Tyr Gly Ser Tyr Ala Lys Pro Thr Asn Glu Gln Gly Gly Gln Gly Thr	725	730	735	
aat act gta gat ctg cag ttc ttt gcc tct tca tgc gct acc tct acg				18988
Asn Thr Val Asp Leu Gln Phe Phe Ala Ser Ser Ser Ala Thr Ser Thr	740	745	750	
cct aaa gcc gta ctc tat gcc gag gac gtg gca ata gaa gca cca gac				19036
Pro Lys Ala Val Leu Tyr Ala Glu Asp Val Ala Ile Glu Ala Pro Asp	755	760	765	
acc cat ttg gtg tac aaa ccg gca gtt aca acc acg acc act agt tcc				19084
Thr His Leu Val Tyr Lys Pro Ala Val Thr Thr Thr Thr Ser Ser	770	775	780	

caa gac ctg cta act cag cag gct gct ccc aac cga ccc aac tac att Gln Asp Leu Leu Thr Gln Gln Ala Ala Pro Asn Arg Pro Asn Tyr Ile 785 790 795 800	19132
ggc ttc agg gat aat ttt atc ggt ctc atg tat tac aac tcc act ggc Gly Phe Arg Asp Asn Phe Ile Gly Leu Met Tyr Tyr Asn Ser Thr Gly 805 810 815	19180
aat atg ggt gtt ttg gca ggg caa gct tct cag cta aac gcg gtg gtt Asn Met Gly Val Leu Ala Gly Gln Ala Ser Gln Leu Asn Ala Val Val 820 825 830	19228
gac ttg caa gac aga aac acc gag ctg tcc tac cag ctc atg ctt gat Asp Leu Gln Asp Arg Asn Thr Glu Leu Ser Tyr Gln Leu Met Leu Asp 835 840 845	19276
gct ttg ggc gac cgc agt cgt tac ttc tcc atg tgg aac cag gcc gta Ala Leu Gly Asp Arg Ser Arg Tyr Phe Ser Met Trp Asn Gln Ala Val 850 855 860	19324
gac agc tat gac cct gat gtc aga att att gaa aat cat ggt gtg gag Asp Ser Tyr Asp Pro Asp Val Arg Ile Ile Glu Asn His Gly Val Glu 865 870 875 880	19372
gat gag ctg cca aac tac tgt ttc ccg cta gga ggg tcg cta gta act Asp Glu Leu Pro Asn Tyr Cys Phe Pro Leu Gly Gly Ser Leu Val Thr 885 890 895	19420
gaa act tat aca ggc cta tca ccc caa aac gga agt aac acg tgg aca Glu Thr Tyr Thr Gly Leu Ser Pro Gln Asn Gly Ser Asn Thr Trp Thr 900 905 910	19468
acc gac agc acc acc tat gca act aga ggg gtg gaa atc ggc tct ggc Thr Asp Ser Thr Thr Tyr Ala Thr Arg Gly Val Glu Ile Gly Ser Gly 915 920 925	19516
aac atg ttc gcc atg gaa att aat ttg gcg gcc aat cta tgg agg agt Asn Met Phe Ala Met Glu Ile Asn Leu Ala Ala Asn Leu Trp Arg Ser 930 935 940	19564
ttc ctg tac tcc aac gtg gcc ctg tac ctg ccc gac gag tac aag ctc Phe Leu Tyr Ser Asn Val Ala Leu Tyr Leu Pro Asp Glu Tyr Lys Leu 945 950 955 960	19612
acc ccc gac aac atc acc ctc ccc gag aac aaa aac act tac gac tac Thr Pro Asp Asn Ile Thr Leu Pro Asp Asn Lys Asn Thr Tyr Asp Tyr 965 970 975	19660
atg aac ggc cgc gtg gcc gcc ccc agc tcc ctc gac acc tac gtc aac Met Asn Gly Arg Val Ala Ala Pro Ser Ser Leu Asp Thr Tyr Val Asn 980 985 990	19708
atc ggg gcg cgc tgg tcc ccc gac ccc atg gac aac gtc aac ccc ttc Ile Gly Ala Arg Trp Ser Pro Asp Pro Met Asp Asn Val Asn Pro Phe 995 1000 1005	19756

aac cac cac cgc aac gcg gga Asn His His Arg Asn Ala Gly 1010 1015	ctg cgc tac cgc tcc atg ctg ctg Leu Arg Tyr Arg Ser Met Leu Leu 1020	19801
ggc aac ggc cgc tac gta ccc Gly Asn Gly Arg Tyr Val Pro 1025 1030	ttc cac atc caa gtg ccc cag aaa Phe His Ile Gln Val Pro Gln Lys 1035	19846
ttc ttc gcc atc aaa aac ctc Phe Phe Ala Ile Lys Asn Leu 1040 1045	ctg ctc ctc ccc ggg tcc tac acc Leu Leu Leu Pro Gly Ser Tyr Thr 1050	19891
tac gag tgg aac ttc cgc aag Tyr Glu Trp Asn Phe Arg Lys 1055 1060	gac gtc aac atg atc ctc cag agc Asp Val Asn Met Ile Leu Gln Ser 1065	19936
agc ctg ggt aac gac ctc cgc Ser Leu Gly Asn Asp Leu Arg 1070 1075	gtc gac ggg gcc agc gtc agg ttc Val Asp Gly Ala Ser Val Arg Phe 1080	19981
gac agc atc aac ctg tac gcc Asp Ser Ile Asn Leu Tyr Ala 1085 1090	aac ttc ttc ccc atg gcc cac aac Asn Phe Phe Pro Met Ala His Asn 1095	20026
acc gcc tcc acg ctc gag gcc Thr Ala Ser Thr Leu Glu Ala 1100 1105	atg ctg cgc aac gac acc aac gac Met Leu Arg Asn Asp Thr Asn Asp 1110	20071
cag tcg ttc aac gac tac ctc Gln Ser Phe Asn Asp Tyr Leu 1115 1120	tgc gct gcc aac atg ctc tac ccc Cys Ala Ala Asn Met Leu Tyr Pro 1125	20116
atc ccc gcc aac gcc acc agc Ile Pro Ala Asn Ala Thr Ser 1130 1135	gtg ccc atc tcc att ccc tcg cgg Val Pro Ile Ser Ile Pro Ser Arg 1140	20161
aac tgg gcc gcc ttc cgg gcc Asn Trp Ala Ala Phe Arg Gly 1145 1150	tgg agc ttc acc cgg ctc aag acc Trp Ser Phe Thr Arg Leu Lys Thr 1155	20206
aag gag acc ccc tct ctg gcc Lys Glu Thr Pro Ser Leu Gly 1160 1165	tcc gcc ttc gat ccc tac ttc acc Ser Gly Phe Asp Pro Tyr Phe Thr 1170	20251
tac tcg ggc tcc atc ccc tac Tyr Ser Gly Ser Ile Pro Tyr 1175 1180	ctg gac gcc acc ttc tac ctc aac Leu Asp Gly Thr Phe Tyr Leu Asn 1185	20296
cac act ttc aag aag gtc tcc His Thr Phe Lys Lys Val Ser 1190 1195	atc atg ttc gac tcc tcc gtc agc Ile Met Phe Asp Ser Ser Val Ser 1200	20341
tgg ccc ggc aac gac cgc ctg Trp Pro Gly Asn Asp Arg Leu 1205 1210	ctg acc ccc aac gag ttc gag atc Leu Thr Pro Asn Glu Phe Glu Ile 1215	20386
aag cgc acc gtg gac ggg gaa	ggg tac aac gtg gcc cag tgc aac	20431

Lys	Arg	Thr	Val	Asp	Gly	Glu	Gly	Tyr	Asn	Val	Ala	Gln	Cys	Asn	
1220						1225					1230				
atg	acc	aag	gac	tgg	ttc	ctc	atc	cag	atg	ctc	agc	cac	tac	aac	20476
Met	Thr	Lys	Asp	Trp	Phe	Leu	Ile	Gln	Met	Leu	Ser	His	Tyr	Asn	
1235						1240					1245				
atc	ggc	tac	cag	ggc	ttc	tac	gtg	ccc	gag	ggc	tac	aag	gac	agg	20521
Ile	Gly	Tyr	Gln	Gly	Phe	Tyr	Val	Pro	Glu	Gly	Tyr	Lys	Asp	Arg	
1250						1255					1260				
atg	tac	tct	ttc	ttc	cgc	aac	ttc	caa	ccc	atg	agc	cgc	cag	gtg	20566
Met	Tyr	Ser	Phe	Phe	Arg	Asn	Phe	Gln	Pro	Met	Ser	Arg	Gln	Val	
1265						1270					1275				
gtc	gac	acc	acc	acc	tac	acc	gac	tac	aaa	aac	gtc	acc	ctc	ccc	20611
Val	Asp	Thr	Thr	Thr	Tyr	Thr	Asp	Tyr	Lys	Asn	Val	Thr	Leu	Pro	
1280						1285					1290				
ttc	cag	cac	aac	aac	tcg	ggg	ttc	gtg	gga	tac	atg	ggc	ccc	acc	20656
Phe	Gln	His	Asn	Asn	Ser	Gly	Phe	Val	Gly	Tyr	Met	Gly	Pro	Thr	
1295						1300					1305				
atg	cgc	gag	ggg	cag	gcc	tac	ccc	gcc	aac	tac	ccc	tac	ccc	ctg	20701
Met	Arg	Glu	Gly	Gln	Ala	Tyr	Pro	Ala	Asn	Tyr	Pro	Tyr	Pro	Leu	
1310						1315					1320				
atc	ggc	aag	acc	gcc	gtg	ccc	agc	ctc	acg	cag	aaa	aag	ttc	ctc	20746
Ile	Gly	Lys	Thr	Ala	Val	Pro	Ser	Leu	Thr	Gln	Lys	Lys	Phe	Leu	
1325						1330					1335				
tgc	gac	cgc	acc	atg	tgg	cgc	atc	ccc	ttc	tcc	agt	aac	ttc	atg	20791
Cys	Asp	Arg	Thr	Met	Trp	Arg	Ile	Pro	Phe	Ser	Ser	Asn	Phe	Met	
1340						1345					1350				
tcc	atg	ggg	gcg	ctc	acc	gac	ctg	ggg	cag	aac	atg	ctg	tac	gcc	20836
Ser	Met	Gly	Ala	Leu	Thr	Asp	Leu	Gly	Gln	Asn	Met	Leu	Tyr	Ala	
1355						1360					1365				
aac	tcc	gcc	cac	gcc	ctc	gac	atg	acc	ttc	gag	gtg	gac	ccc	atg	20881
Asn	Ser	Ala	His	Ala	Leu	Asp	Met	Thr	Phe	Glu	Val	Asp	Pro	Met	
1370						1375					1380				
gat	gag	ccc	acg	ctt	ctc	tat	gtt	ctg	ttc	gaa	gtg	ttc	gac	gtc	20926
Asp	Glu	Pro	Thr	Leu	Leu	Tyr	Val	Leu	Phe	Glu	Val	Phe	Asp	Val	
1385						1390					1395				
gtg	cgc	atc	cac	cag	ccg	cac	cgc	ggc	gtc	atc	gag	gcc	gtc	tac	20971
Val	Arg	Ile	His	Gln	Pro	His	Arg	Gly	Val	Ile	Glu	Ala	Val	Tyr	
1400						1405					1410				
ctg	cgc	acg	ccg	ttc	tcg	gcc	ggg	aac	gcc	acc	acc	taa	ggagggggcc	21020	
Leu	Arg	Thr	Pro	Phe	Ser	Ala	Gly	Asn	Ala	Thr	Thr				
1415						1420					1425				
gccgacggat	gggctccagc	gagccggagc	tggtcgccat	cgcgcgcgac	ctgggctgcg	21080									

ggccctactt cctgggcacc tttagaaaac gcttcccggg ctctgtggcg ccgcacaagc 21140
 tggcctgcgc catcgtcaac accgcgggac gcgagaccgg cggcgtccac tggctggccc 21200
 tggcctggaa ccccccgcagc cgaacctgct acctcttcga ccccttcggc ttctcggacg 21260
 acaggctcag gcagatctac cagttcgagt acgaaggcct gctccggcgc agcgccctcg 21320
 cctccacccc cgaccactgc gtcaccctcg tcaagtccac ccagaccgtc cagggggccc 21380
 gctcggcgcg ctgcggcctc ttctgtgtga tgttctgtga cgccttcgtg cgctggcccg 21440
 cctcccccat ggacggcaac cccaccatgg acctccttac gggcgttccc aacagcatgc 21500
 ttcagagtcc ccaggctcag cccaccctcc accgcaacca ggaggaactc tacgccttcc 21560
 tggctcggca ctccccctac ttctgcgcgc accgcgagcg catagaaaag gccaccgagt 21620
 ttgacaaaat gaacgactag attttctgtg aaaaacactc aataaagcct ttattgggtc 21680
 accacacgtg cacgcagtga gactttttat taaaagggc tccgcctcct cgtcgcctg 21740
 gctggtgggg agggagacgt tgcgatactg caggcgggag ctccatctga actcgggaat 21800
 cagcagcttg ggcagggggc cctcgacgtt ctgcctccac agcttgcgca ccagctgcag 21860
 ggcgccagc aggtcggggc cggagatcct gaagtgcag ttggggccct ggttgccgcg 21920
 ggagttgcgg tacaccgggt tggcgactg gaacaccagc acgctggggg gctcgaatgct 21980
 ggccagcgcc gtcttgctgg tcacctcgtc gccgcgcagg gactccgcgt tgcctcagcg 22040
 gaaggcggtc agcttgacaa gctgcccacc cagcacgggc accccgctcg gctgggtcag 22100
 gcagtcgcag cgcatagcca tcagcagccg cttctgcccg tgctgcatct tcggatagtc 22160
 ggctcgcagt aaggcctcca tctgcccga ggccgtctgc gccttgctgc cctccgagaa 22220
 gaacagcccc caggacttgc cggagaacac gttgttgccg cagctcacgt cttccacgca 22280
 gcagcgcgcg tcgtcgttct teagctgcac cacgtgcgg cccagcgggt tctgcaccac 22340
 cttggtcttg cgggatgtt ccttcagggc ccgctggccg ttctcgtggt tcacgtccat 22400
 ctccaccacc tgctccttct ggatcatctc cagcccgtgg tagcagcgca gcacgccctc 22460
 ctgctcgggt caccctgca gccagacggc gcagccggtc ggctccagct gttgaggtt 22520
 cccccggcg taggtctcca cgtacggccg caggaagcgg cccatcatct ccacaaaggt 22580
 cttctgaccg gtgaaggtca gctgcagccc gcgatgctcc tcgttgagcc acgtctgaca 22640
 gatcttgccg tacaccttgc cctgctcggg cagaaacttg aaagcggcct tctcctcggg 22700
 ctccacgtgg tacttctcca tcagcgccga catcagctcc atgccttct cccaggccga 22760
 caccagcggc tccgcgcggg ggttcaccac cgccatgcct cgggaagtgc cggggcgctc 22820

atcttctctcc	tcctcctcgt	cttcttcttg	aggcggcggg	ggcggcagtt	gtctcacgaa	22880
tctcttgccg	ttggccttct	ggacgatctc	cacgccgggg	tgggtgaacc	cgtgggccac	22940
caccacttcg	tcctcttctc	cttcgctgtc	gggcacgact	tcgggagagg	gaggcggcgg	23000
aggaaccggg	gcggccactg	cggccatcgc	ggcgttcttg	cgcgccttct	tggggggcag	23060
aggcggcgtc	tcgcgctccg	ggctggtctc	ttgcaggtag	ggcgtgatgg	tgtgggaggt	23120
ggggcgctct	ggctgacggc	cggccatgct	gatgcttgac	tcctaggcga	aaagatggag	23180
gaggatctta	gacagccgca	gccctgtctc	gaaaccttaa	ccacccccgc	ctctgaggtc	23240
ggcgcggggc	agctagacat	gcaacgggag	gaggaggagg	acgtgcgagt	ggagcaagac	23300
ccgggctacg	tgacgccgcc	cgaggacggc	gaggagccgc	aggcaccggc	gccaacgctc	23360
agcgaagccg	actacctggg	aggggaggac	gacgtgctgc	tgaagcacct	ggcgcggcag	23420
agcaccatcg	tgaggagggc	cctcaaggag	cgcgaggagg	tcccgcgtgac	ggtggaggag	23480
ctcagccggg	cctacgaagc	caacctcttc	tcgccgcggg	tgccccccaa	gaagcaggcc	23540
aacggcacct	gcgagcccaa	ccccgcctc	aacttctacc	ccgtctttgc	ggtgcccgag	23600
gcgctggcca	cctatcacat	cttcttcaag	aaccagcgca	tccccctctc	gtgccgcgcc	23660
aaccgcaccc	gcgccgaccg	cctcctgcat	ctccgagccg	gcgccgccat	acctgagatc	23720
gcctccctgg	aggaagtccc	caagatcttc	gaaggctctg	gcaaggacga	gaagcgcgcg	23780
gcaaacgctc	tggaaaagaa	cgagagcgag	ggtcagaacg	tgctggtcga	gctggaaggc	23840
gacaacgcgc	gtctggccgt	gctcaaacgc	accatcgaag	tctcccactt	cgcctacccc	23900
gcgctcaacc	ttccccccaa	ggcatgcgc	tcggatcatg	atcagctgct	catcaagcgc	23960
gccgagcccc	tcgagaacga	ctccgagggt	gattccgagg	acggaaaacc	cgtgggtctg	24020
gacgaggagc	tcgcgcgctg	gctgggcacg	caggaccccc	ccgagttgca	agagcggcgc	24080
aagatgatga	tggcggccgt	gctggtcacc	gccgagctcg	agtgcctgca	gcgcttcttc	24140
gccgaccccc	agaccctgcg	caaggctcag	gagtcctctc	actacgcctt	ccgccacggc	24200
tacgtgcgcc	aggcctgcaa	gatctccaac	gtggagctta	gcaacctggg	ctcctacatg	24260
ggcatcctgc	acgagaaccg	cctcgggcag	aacgtcctcc	actgcaccct	gaccggggag	24320
gcccgcgcgc	actacgtccg	cgactgcate	tacctctttc	tcaccctcac	ctggcagacc	24380
gccatggggg	tctggcagca	gtgtctggag	gagcgcaacc	ttcgcgagct	cgacaagcta	24440
ctgagccgcg	agcgcgcgca	gctctggacg	gctttcagcg	agcgcaccgc	cgcctgccgt	24500

ctggccgacc tcatcttccc cgagcgactc aggcacaccc tccagaacgg cctgcccagc 24560
 tttgtcagcc agagcatgct gcaaaacttt cgctccttca tcctggagcg atccggcatc 24620
 ttgcccgcga tgagctgcgc cctgccctcc gatttcgtcc cctctatta tcgcgagtgc 24680
 cccccgcgc tctggagcca ctgctacctg ctgcgtctgg ccaactacct cgcccaccac 24740
 tccgacctca tggaagactc cagcggcgag gggctgctgg agtgccactg ccgctgcaac 24800
 ctctgcaccc cccaccgctc gctggtctgc aacaccgagc tgctcagcga gacgcaagtg 24860
 atcggtagct ttgagatcca gggactagag gggccggagg gtgcttccaa cctcaagctc 24920
 agccccgcgc tctggacttc cgctacctg cgcaaattta tccccgagga ctatcacgcc 24980
 caccagatcc aattctacga agaccaatcg cgacccccca aagccccct cacggcctgt 25040
 gtcacacccc agagccagat tctggcccaa ttgcaagcca tccagcaggc ccgccaagag 25100
 ttcctcctga aaaaggggtca cgggggtctat ctggaccccc agaccggcga ggaactcaac 25160
 acccgcgcac cctccgccc cgcttcgtgc cgcccgaga accatgccgc ccaaagggaa 25220
 caagcaggcc atcgcccagc ggcggggccaa gaagcagcaa gagctccagg agcagtggga 25280
 cgaggagtcc tgggacagcc aggcggagga agtctcagac gaggaggagg acatggagag 25340
 ctgggacagc ctagacgagg agggaggagg cgaggagcta gaggacgagc ctctcgagga 25400
 ggaagagccc agcagcgccg cggcaccatc ggcttccaaa gaagcggctc ggagccggcc 25460
 ggccccgaag cagcagaagc agcaacagcc gccaccgtcg ccccgacgc caccaccagg 25520
 ctactcaaa gccagccgta ggtgggacgc ggtgtccatc gcgggatcgc ccaaagcccc 25580
 agtcggttaag ccacccgggc ggtcgccggc ggggtactgt tcctggcgcc ccacaagag 25640
 caagatcgtc gcctgcctcc agcactgccg gggcaacatc tccttcgcgc ggcgctaact 25700
 gctcttccac gacgggggtg cgggtgcgcg caacgtctc tactattacc gteatctcta 25760
 cagcccctac gagacagaag gcccggcctc cgcgtaagac cagccgccag acgggtctct 25820
 ccgccatcgc gacccgccag gactcggccg ccacgcagga gtcagaaaa cgcactcttc 25880
 ccaccctgta tgctatcttc cagcagagcc gcggccagca gctggaactg aaagtaaaaa 25940
 accgctccct gcgttcgctc acccgagct gtctgtacca caggagggaa gaccaactgc 26000
 agcgcacgct cgaggacgcc gaggcactgt tcaataaata ctgctcgggt tctcttaagg 26060
 actgaaagcc cgcgcttttt cagaggctca ttacgtcatc atcatcatga gcaaggacat 26120
 tcccacgcct tacatgtgga gctaccagcc gcagatggga ctggcggccg gcgcctccca 26180
 ggattactcc agtcgcatga actggtgag tgccggcccc cacatgatcg ggcgggtcaa 26240

tgggattcgt gccacccgca atcagatact gctggaacag gccgccctca cctccacccc	26300
gcgacgtcag ctgaacccgc ccgcttgcc cgccgccag gtgtaccagg aaaacccgc	26360
cccgaccaca gtcctcctgc cagcgacgc ggagggcgaa gtccagatga ctaactccgg	26420
ggcgcaatta gcggggcgccg ccgccacgt cgtcgctccc gggtacagag gtcggcccg	26480
accctacccc tccggcccta taagaggct gatcattcga ggccgaggta tccagctcaa	26540
cgacgagggtg gtgagctcct cgaccggtct tcggcccgac ggagtcttcc agcttgagg	26600
cgccggccgc tcttccttca cactcgcca ggctacctg acgctccaga gctcttctc	26660
ccagcctcgc tccggcgga tcggcaccct ccagttcgtg gaggagttcg tgccctcggt	26720
ctacttcaac ccgttctccg gctctcccg ccgtacctg gacagcttca tcccaacta	26780
cgacgcggtg agcgaatccg tggacggcta cgattgatga ccgatgggtgc ggccgtaact	26840
gcggggcggc aacatctgca tcaactgcat cgtcctcggt gcttcgccc ggaggcctgt	26900
gagttcatct acttccagct cggcccgac cagcttcagg gcccttcgca cggcgtaag	26960
ctcgtgatag aggaagagct cgagagtagc tgctgcgct gttttacctc gcgccccatc	27020
ctagtcgaga gggaacgcgg taggaccacc ctcacctct actgcattcg tgactccccg	27080
gaattacatg aagatctgtg ttgccttcta tgtgccgaac aataaccct cttgtaacta	27140
cctacatcca caataaacca gaatttgaa actccttctg tttgtttgca g atg aaa	27197
	Met Lys
cgc gcc cgc ctc gac gac gac ttc aac ccc gtc tac ccc tat gac	27242
Arg Ala Arg Leu Asp Asp Asp Phe Asn Pro Val Tyr Pro Tyr Asp	
1430 1435 1440	
act ccc aac gct ccg tct gtt ccc ttc atc act cct ccc ttc gtc	27287
Thr Pro Asn Ala Pro Ser Val Pro Phe Ile Thr Pro Pro Phe Val	
1445 1450 1455	
tcc tcg gac ggc ttg caa gaa aaa cca ccc gga atg ctc agt ctc	27332
Ser Ser Asp Gly Leu Gln Glu Lys Pro Pro Gly Met Leu Ser Leu	
1460 1465 1470	
aac tac caa gat cct att acc acc caa aac ggg gca tta act cta	27377
Asn Tyr Gln Asp Pro Ile Thr Thr Gln Asn Gly Ala Leu Thr Leu	
1475 1480 1485	
aag ctt ggc agc gga ctg aac ata aac caa gat ggg gaa ctt acc	27422
Lys Leu Gly Ser Gly Leu Asn Ile Asn Gln Asp Gly Glu Leu Thr	
1490 1495 1500	
tca gac gcc agc gtt ctc gtc act ccc ccc att aca aaa gcc aac	27467
Ser Asp Ala Ser Val Leu Val Thr Pro Pro Ile Thr Lys Ala Asn	

1505				1510				1515				
aac aca ata	ggc cta gcc ttc	aat gca cct ctt acc ttg	caa agc	27512								
Asn Thr Ile	Gly Leu Ala Phe	Asn Ala Pro Leu Thr Leu	Gln Ser									
1520		1525	1530									
gat act tta	aat ctt gct tgt aac	gcc cca ctt acc gtg	caa gac	27557								
Asp Thr Leu	Asn Leu Ala Cys	Asn Ala Pro Leu Thr Val	Gln Asp									
1535		1540	1545									
aat agg ttg	gga ata aca tac aac	tct ccc ctc acc ttg	caa aac	27602								
Asn Arg Leu	Gly Ile Thr Tyr	Asn Ser Pro Leu Thr Leu	Gln Asn									
1550		1555	1560									
agc gaa ctt	gcc cta gcg gtc acc	ccg cct ctt gac act	gcc aat	27647								
Ser Glu Leu	Ala Leu Ala Val Thr	Pro Pro Leu Asp Thr	Ala Asn									
1565		1570	1575									
aac aca ctt	gcg ctt aaa acc gcc	cgg cct ata att aca	aac tct	27692								
Asn Thr Leu	Ala Leu Lys Thr	Arg Pro Ile Ile Thr	Asn Ser									
1580		1585	1590									
aat aac gag	ctt aca ctg tdc gct	gat gct ccc cta aac	acc agc	27737								
Asn Asn Glu	Leu Thr Leu Ser	Ala Asp Ala Pro Leu Asn	Thr Ser									
1595		1600	1605									
acg ggt acc	ctc cgc cta caa agc	gca gca cca ctg ggg	cta gtt	27782								
Thr Gly Thr	Leu Arg Leu Gln	Ala Ala Pro Leu Gly	Leu Val									
1610		1615	1620									
gac caa acc	ctg cga gtg ctt ttt	tct aac cca ctc tac	ttg caa	27827								
Asp Gln Thr	Leu Arg Val Leu Phe	Ser Asn Pro Leu Tyr	Leu Gln									
1625		1630	1635									
aac aac ttt	ctc tca cta gcc att	gaa cgc cca ttg gct	tta act	27872								
Asn Asn Phe	Leu Ser Leu Ala Ile	Glu Arg Pro Leu Ala	Leu Thr									
1640		1645	1650									
acc act ggt	tct atg gct atg cag	att tcc caa cca tta	aaa gtg	27917								
Thr Thr Gly	Ser Met Ala Met Gln	Ile Ser Gln Pro Leu	Lys Val									
1655		1660	1665									
gaa gac gga	agc tta agc ttg agc	att gaa agc cct cta	aat cta	27962								
Glu Asp Gly	Ser Leu Ser Leu Ser	Ile Glu Ser Pro Leu	Asn Leu									
1670		1675	1680									
aaa aac gga	aat ctt act tta gga	acc caa agt ccc cta	act gtc	28007								
Lys Asn Gly	Asn Leu Thr Leu Gly	Thr Gln Ser Pro Leu	Thr Val									
1685		1690	1695									
act ggt aac	aac ctc agc ctt aca	aca aca gcc cca tta	acg gtt	28052								
Thr Gly Asn	Asn Leu Ser Leu Thr	Thr Thr Ala Pro Leu	Thr Val									
1700		1705	1710									
cag aac aac	gct cta gcc ctc tca	gtg tta ctg ccg ctt	aga cta	28097								
Gln Asn Asn	Ala Leu Ala Leu Ser	Val Leu Leu Pro Leu	Arg Leu									
1715		1720	1725									

ttt aat aac acc tca ctg gga gtg gca ttc aac cca ccc att tct Phe Asn Asn Thr Ser Leu Gly Val Ala Phe Asn Pro Pro Ile Ser 1730 1735 1740	28142
tca gca aac aac ggg ctg tct ctt gac att gga aat ggc ctt aca Ser Ala Asn Asn Gly Leu Ser Leu Asp Ile Gly Asn Gly Leu Thr 1745 1750 1755	28187
ctg caa tac aac agg ctc gta gtg aac att ggc ggc ggg cta cag Leu Gln Tyr Asn Arg Leu Val Val Asn Ile Gly Gly Gly Leu Gln 1760 1765 1770	28232
ttt aac aac ggt gct att acc gct tcc ata aat gca gct ctg ccg Phe Asn Asn Gly Ala Ile Thr Ala Ser Ile Asn Ala Ala Leu Pro 1775 1780 1785	28277
ttg cag tat tcc aat aac cag ctt tct ctt aat att gga ggc ggg Leu Gln Tyr Ser Asn Asn Gln Leu Ser Leu Asn Ile Gly Gly Gly 1790 1795 1800	28322
ctg cga tac aac ggc act tac aaa aat tta gcc gtc aaa acc gac Leu Arg Tyr Asn Gly Thr Tyr Lys Asn Leu Ala Val Lys Thr Asp 1805 1810 1815	28367
tct ttt agg ggt ctt gaa att gac agt aat cag ttc ctg gtg cca Ser Phe Arg Gly Leu Glu Ile Asp Ser Asn Gln Phe Leu Val Pro 1820 1825 1830	28412
aga ctg ggt tct ggt cta aag ttt gat caa tat ggg tac att agc Arg Leu Gly Ser Gly Leu Lys Phe Asp Gln Tyr Gly Tyr Ile Ser 1835 1840 1845	28457
gtc ata cct cca act gtt acg cca aca aca ctt tgg act aca gca Val Ile Pro Pro Thr Val Thr Pro Thr Thr Leu Trp Thr Thr Ala 1850 1855 1860	28502
gac cct tct ccc aac gct act ttt tac gac agc tta gat gct aag Asp Pro Ser Pro Asn Ala Thr Phe Tyr Asp Ser Leu Asp Ala Lys 1865 1870 1875	28547
gta tgg ctg gcc tta gta aaa tgc aac ggc atg gtt aat gga acc Val Trp Leu Ala Leu Val Lys Cys Asn Gly Met Val Asn Gly Thr 1880 1885 1890	28592
ata gcc ata aag gct tta aaa ggt act ctg ctc caa cct acg gct Ile Ala Ile Lys Ala Leu Lys Gly Thr Leu Leu Gln Pro Thr Ala 1895 1900 1905	28637
agt ttt att tct ttt gtt atg tat ttt tac agc aat ggc acc aga Ser Phe Ile Ser Phe Val Met Tyr Phe Tyr Ser Asn Gly Thr Arg 1910 1915 1920	28682
aga act aac tac ccc acg ttt gaa aat gaa ggc ata cta gct agt Arg Thr Asn Tyr Pro Thr Phe Glu Asn Glu Gly Ile Leu Ala Ser 1925 1930 1935	28727

agt gct aca tgg ggt tat cgt caa gga aac tcg gca aac acc aac	28772
Ser Ala Thr Trp Gly Tyr Arg Gln Gly Asn Ser Ala Asn Thr Asn	
1940 1945 1950	
gtc acc agt gcc gtt gaa ttt atg cct agc tcc aca aga tat cct	28817
Val Thr Ser Ala Val Glu Phe Met Pro Ser Ser Thr Arg Tyr Pro	
1955 1960 1965	
gtt aac aag ggt act gag gtt cag aac atg gaa ctc acc tac act	28862
Val Asn Lys Gly Thr Glu Val Gln Asn Met Glu Leu Thr Tyr Thr	
1970 1975 1980	
ttc ttg cag gga gac ccc act atg gcc ata tca ttt caa gct att	28907
Phe Leu Gln Gly Asp Pro Thr Met Ala Ile Ser Phe Gln Ala Ile	
1985 1990 1995	
tat aac cat gct ttg gaa ggt tac tct tta aaa ttt acc tgg cga	28952
Tyr Asn His Ala Leu Glu Gly Tyr Ser Leu Lys Phe Thr Trp Arg	
2000 2005 2010	
gtt cgc aac agg gaa cgc ttt gat atc ccc tgc tgt tct ttt tct	28997
Val Arg Asn Arg Glu Arg Phe Asp Ile Pro Cys Cys Ser Phe Ser	
2015 2020 2025	
tac ata acg gaa gaa taa acactgtttt tcttttcaat gtttttatcc	29045
Tyr Ile Thr Glu Glu	
2030	
tgctttttta cacagttcga accgtcagac tccctccccc cttccacttc acccggtaca	29105
cctcccgcgc cccctggatc gctgcgtaca actgcagttt ggtgttcaga cacgggttct	29165
taggtgacag tatecacacg gcctctttgc cggccaggcg ctggtccgta atgctcacia	29225
atccctccga cagctcctcc agacacacgg tggaatccaa ggccgcccgc taaaaacaa	29285
acacagtcac gctctccacg ggttctctcc tgggtcgtac tgcgccagcg tgaacgggcg	29345
atggtgctcc atcagggctc gcagcaaccg ctgtcggcgc ggctcaccga ggctcggcg	29405
aaaagcgcgc cgtctgggag tgctattcaa aaaacgcacc gcctttatca acagtctcct	29465
cgtgcggcgg gcgcagcagc gcacctggat ctctgtcagg tctttacaat aggtacagcc	29525
catcaccacc atgttggtta aaatcccaaa gctaaacacg ctccacccaa atgacatgaa	29585
ttccagcacc gccgcggcgt ggccatcata caatatgcgg aggtaaatca ggtgccgccc	29645
cctaatacaa acgctcccca tatacatcac ctcttaggc agttgataat taaccacctc	29705
ccggtaccag ggaaacctca cgtttactaa agcccaaac accaacattt taaaccagtt	29765
agccagcacc acccctcccg ccttactcag cagcgacccc ggctgtttac aatgacagtg	29825
aatcaccacac ctctcatacc ccctaataac ctggcgtggc tccacatcta tagtagcaca	29885
gcacacgcac accctcatgt aatgcttcat cacaaatctt tcccaagggg ttagtatcat	29945

gtcccagggg	acggggccact	cctgcagcac	gggtgaaaggt	acgcaggcgg	gaacagtcc	30005
cacctcggac	acataatgca	tattcagatg	ttcacactct	aaaaccccgg	ggcttccctc	30065
caacgcagcc	actggcaagt	tctcagaggg	tggtgtaagg	cgggtggtgct	gatagggact	30125
caatctgtgt	cgacaccgtc	tgtcgcgttg	catcgtagac	caacgcttgg	cgacaccgct	30185
cgtacttcgc	ccaaagaaaa	cgggtgcgac	gccaacacac	ttccgcgtac	cgtgggttcc	30245
gcactcgagc	tcgctcagtt	ctcaacgcat	aatgcagcca	ttcctgtaat	ccacacaaca	30305
gtcgctcggc	ttccaaagag	atgtgcacct	cgtatcttat	aacgtcccga	tatatatcca	30365
agcaggcagt	cagggccact	tgcaaccagt	gcacgcaggc	ggactgatcg	cgacacactg	30425
gagggtggagg	gagagacgga	agaggcatgt	tactccagac	ggtcgaaaag	cggatcaaag	30485
tgcagatcgc	gaagatggca	gcgatccccg	ccgctacgct	ggatagat	cacagccagg	30545
tcaaacataa	tgcggttttc	caaatacct	attaccgcct	ccaccagagc	cgccacgcgc	30605
acttccagaa	acaccagcac	ggctacggca	ttctcctcaa	aatcttcaaa	cattaagctg	30665
catgattgaa	tcacccccaa	ataattctcc	tccttccatt	ctcgcaaaat	ttgagtaaaa	30725
acctctcgca	gattagctcc	gtggcggttca	aaaagggtcac	ttagagcgcc	ctccaccgcc	30785
atgcgcaagc	acaccctcat	gattgaaaaa	tgccagtctc	ctgaaccacc	tgcaattgat	30845
ttaaaagacc	tatattagga	tcaattccac	tctcccgag	ctccacgcgt	agcattagct	30905
gcaaaaagtc	atttaaattct	tcgcaaacta	gcgcggtaag	ctcgccgccg	ggaattagggt	30965
ctgaagcagt	caccacacac	ataatttcca	gtgaaggagt	cagtctaagc	agcaaaaagc	31025
cgcagtagca	gtgttgaaaa	ggaggggtca	cgcaatgtaa	catatgcagc	caaaaatctc	31085
caagggtgtct	gtgcataaac	tccaccactg	aaaagtccaa	atcatgtaaa	tatgccatca	31145
ccgcctcagg	aaccaccacg	gacacaaaaa	cgggccgtag	caaatacatg	gtgtcctgca	31205
aagcaaaaac	acatttatac	catagaggcg	cgaattactt	ggggaaaaat	cactcgctcc	31265
aaaactaaac	aggccaccgt	ctgaccgcgc	cagccataaa	aaaagcgggt	cgaatgatta	31325
aaaagaataa	tagacacctc	ccaccaggta	ctcggtgca	actcgtcgc	ccctatcaaa	31385
accccgcgga	cgttcagtgc	ggccatagaa	aaaatgcggc	ccaaatatcc	caccggaatc	31445
tccacggcca	gctgcagtga	tagcaaaaga	acgccatgag	gagcaatcac	aaaattttca	31505
ggcgataaaa	gcacataaag	gttagaatag	ccctgctgca	caggtaataa	agcccgcgag	31565
ctcagcaaat	gcacataaac	cgttcagcc	atcccgctct	accgcgaaca	aaaggctcac	31625

agtacacagt tactcaaccc acacgccaca cagtatttat acactcctca atcgccacgt 31685
 caccgcgcc gaacaaactc caaaagtcca aaaagtccaa aacgcccgcg taaaagcccg 31745
 ccaaaacagc acttcctcat ttactctccc acagtacgtc acttccgccc cgcccgcgcg 31805
 cctcgccccg cctcaccct cgcgtccac cccgcgcccc acgtcagact cccacccgcg 31865
 cccgcgcccc cgtcatccgc acccaccct cactccaccc ctaaccccg ctcctcatta 31925
 tcatattggc accgttccca aataagggtat attatgatga tg 31967

<210> 13
 <211> 508
 <212> PRT
 <213> Simian adenovirus

<400> 13

Met Arg Arg Ala Val Gly Val Pro Pro Val Met Ala Tyr Ala Glu Gly
 1 5 10 15

Pro Pro Pro Ser Tyr Glu Thr Val Met Gly Ala Ala Asp Ser Pro Ala
 20 25 30

Thr Leu Glu Ala Leu Tyr Val Pro Pro Arg Tyr Leu Gly Pro Thr Glu
 35 40 45

Gly Arg Asn Ser Ile Arg Tyr Ser Glu Leu Ala Pro Leu Tyr Asp Thr
 50 55 60

Thr Arg Val Tyr Leu Val Asp Asn Lys Ser Ala Asp Ile Ala Ser Leu
 65 70 75 80

Asn Tyr Gln Asn Asp His Ser Asn Phe Leu Thr Thr Val Val Gln Asn
 85 90 95

Asn Asp Phe Thr Pro Val Glu Ala Gly Thr Gln Thr Ile Asn Phe Asp
 100 105 110

Glu Arg Ser Arg Trp Gly Gly Asp Leu Lys Thr Ile Leu Arg Thr Asn
 115 120 125

Met Pro Asn Ile Asn Glu Phe Met Ser Thr Asn Lys Phe Arg Ala Arg
 130 135 140

Leu Met Val Glu Lys Val Asn Lys Glu Thr Asn Ala Pro Arg Tyr Glu

145 150 155 160
 Trp Phe Glu Phe Thr Leu Pro Glu Gly Asn Tyr Ser Glu Thr Met Thr
 165 170 175
 Ile Asp Leu Met Asn Asn Ala Ile Val Asp Asn Tyr Leu Glu Val Gly
 180 185 190
 Arg Gln Asn Gly Val Leu Glu Ser Asp Ile Gly Val Lys Phe Asp Thr
 195 200 205
 Arg Asn Phe Arg Leu Gly Trp Asp Pro Val Thr Lys Leu Val Met Pro
 210 215 220
 Gly Val Tyr Thr Asn Glu Ala Phe His Pro Asp Ile Val Leu Leu Pro
 225 230 235 240
 Gly Cys Gly Val Asp Phe Thr Gln Ser Arg Leu Ser Asn Leu Leu Gly
 245 250 255
 Ile Arg Lys Arg Met Pro Phe Gln Ala Gly Phe Gln Ile Met Tyr Glu
 260 265 270
 Asp Leu Glu Gly Gly Asn Ile Pro Ala Leu Leu Asp Val Ala Lys Tyr
 275 280 285
 Glu Ala Ser Ile Gln Lys Ala Arg Glu Gln Gly Gln Glu Ile Arg Gly
 290 295 300
 Asp Asn Phe Thr Val Ile Pro Arg Asp Val Glu Ile Val Pro Val Glu
 305 310 315 320
 Lys Asp Ser Lys Asp Arg Ser Tyr Asn Leu Leu Pro Gly Asp Gln Thr
 325 330 335
 Asn Thr Ala Tyr Arg Ser Trp Phe Leu Ala Tyr Asn Tyr Gly Asp Pro
 340 345 350
 Glu Lys Gly Val Arg Ser Trp Thr Leu Leu Thr Thr Thr Asp Val Thr
 355 360 365
 Cys Gly Ser Gln Gln Val Tyr Trp Ser Leu Pro Asp Met Met Gln Asp
 370 375 380

Pro Val Thr Phe Arg Pro Ser Ser Gln Val Ser Asn Tyr Pro Val Val
 385 390 395 400

Gly Val Glu Leu Leu Pro Val His Ala Lys Ser Phe Tyr Asn Glu Gln
 405 410 415

Ala Val Tyr Ser Gln Leu Ile Arg Gln Ser Thr Ala Leu Thr His Val
 420 425 430

Phe Asn Arg Phe Pro Glu Asn Gln Ile Leu Val Arg Pro Pro Ala Pro
 435 440 445

Thr Ile Thr Thr Val Ser Glu Asn Val Pro Ala Leu Thr Asp His Gly
 450 455 460

Thr Leu Pro Leu Arg Ser Ser Ile Ser Gly Val Gln Arg Val Thr Ile
 465 470 475 480

Thr Asp Ala Arg Arg Arg Thr Cys Pro Tyr Val His Lys Ala Leu Gly
 485 490 495

Ile Val Ala Pro Lys Val Leu Ser Ser Arg Thr Phe
 500 505

<210> 14
 <211> 917
 <212> PRT
 <213> Simian adenovirus

<400> 14

Met Ala Thr Pro Ser Met Met Pro Gln Trp Ser Tyr Met His Ile Ala
 1 5 10 15

Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
 20 25 30

Arg Ala Thr Asp Thr Tyr Phe Ser Leu Gly Asn Lys Phe Arg Asn Pro
 35 40 45

Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
 50 55 60

Thr Leu Arg Phe Val Pro Val Asp Arg Glu Asp Thr Ala Tyr Ser Tyr
65 70 75 80

Lys Val Arg Phe Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
85 90 95

Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser
100 105 110

Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
115 120 125

Ala Pro Asn Pro Ser Glu Trp Lys Gly Ser Asp Asn Lys Ile Ser Val
130 135 140

Arg Gly Gln Ala Pro Phe Phe Ser Thr Ser Ile Thr Lys Asp Gly Ile
145 150 155 160

Gln Val Ala Thr Asp Thr Ser Ser Gly Ala Val Tyr Ala Lys Lys Glu
165 170 175

Tyr Gln Pro Glu Pro Gln Val Gly Gln Glu Gln Trp Asn Ser Glu Ala
180 185 190

Ser Asp Ser Asp Lys Val Ala Gly Arg Ile Leu Lys Asp Thr Thr Pro
195 200 205

Met Phe Pro Cys Tyr Gly Ser Tyr Ala Lys Pro Thr Asn Glu Gln Gly
210 215 220

Gly Gln Gly Thr Asn Thr Val Asp Leu Gln Phe Phe Ala Ser Ser Ser
225 230 235 240

Ala Thr Ser Thr Pro Lys Ala Val Leu Tyr Ala Glu Asp Val Ala Ile
245 250 255

Glu Ala Pro Asp Thr His Leu Val Tyr Lys Pro Ala Val Thr Thr Thr
260 265 270

Thr Thr Ser Ser Gln Asp Leu Leu Thr Gln Gln Ala Ala Pro Asn Arg
275 280 285

Pro Asn Tyr Ile Gly Phe Arg Asp Asn Phe Ile Gly Leu Met Tyr Tyr

290

295

300

Asn Ser Thr Gly Asn Met Gly Val Leu Ala Gly Gln Ala Ser Gln Leu
305 310 315 320

Asn Ala Val Val Asp Leu Gln Asp Arg Asn Thr Glu Leu Ser Tyr Gln
325 330 335

Leu Met Leu Asp Ala Leu Gly Asp Arg Ser Arg Tyr Phe Ser Met Trp
340 345 350

Asn Gln Ala Val Asp Ser Tyr Asp Pro Asp Val Arg Ile Ile Glu Asn
355 360 365

His Gly Val Glu Asp Glu Leu Pro Asn Tyr Cys Phe Pro Leu Gly Gly
370 375 380

Ser Leu Val Thr Glu Thr Tyr Thr Gly Leu Ser Pro Gln Asn Gly Ser
385 390 395 400

Asn Thr Trp Thr Thr Asp Ser Thr Thr Tyr Ala Thr Arg Gly Val Glu
405 410 415

Ile Gly Ser Gly Asn Met Phe Ala Met Glu Ile Asn Leu Ala Ala Asn
420 425 430

Leu Trp Arg Ser Phe Leu Tyr Ser Asn Val Ala Leu Tyr Leu Pro Asp
435 440 445

Glu Tyr Lys Leu Thr Pro Asp Asn Ile Thr Leu Pro Asp Asn Lys Asn
450 455 460

Thr Tyr Asp Tyr Met Asn Gly Arg Val Ala Ala Pro Ser Ser Leu Asp
465 470 475 480

Thr Tyr Val Asn Ile Gly Ala Arg Trp Ser Pro Asp Pro Met Asp Asn
485 490 495

Val Asn Pro Phe Asn His His Arg Asn Ala Gly Leu Arg Tyr Arg Ser
500 505 510

Met Leu Leu Gly Asn Gly Arg Tyr Val Pro Phe His Ile Gln Val Pro
515 520 525

Gln Lys Phe Phe Ala Ile Lys Asn Leu Leu Leu Leu Pro Gly Ser Tyr
530 535 540

Thr Tyr Glu Trp Asn Phe Arg Lys Asp Val Asn Met Ile Leu Gln Ser
545 550 555 560

Ser Leu Gly Asn Asp Leu Arg Val Asp Gly Ala Ser Val Arg Phe Asp
565 570 575

Ser Ile Asn Leu Tyr Ala Asn Phe Phe Pro Met Ala His Asn Thr Ala
580 585 590

Ser Thr Leu Glu Ala Met Leu Arg Asn Asp Thr Asn Asp Gln Ser Phe
595 600 605

Asn Asp Tyr Leu Cys Ala Ala Asn Met Leu Tyr Pro Ile Pro Ala Asn
610 615 620

Ala Thr Ser Val Pro Ile Ser Ile Pro Ser Arg Asn Trp Ala Ala Phe
625 630 635 640

Arg Gly Trp Ser Phe Thr Arg Leu Lys Thr Lys Glu Thr Pro Ser Leu
645 650 655

Gly Ser Gly Phe Asp Pro Tyr Phe Thr Tyr Ser Gly Ser Ile Pro Tyr
660 665 670

Leu Asp Gly Thr Phe Tyr Leu Asn His Thr Phe Lys Lys Val Ser Ile
675 680 685

Met Phe Asp Ser Ser Val Ser Trp Pro Gly Asn Asp Arg Leu Leu Thr
690 695 700

Pro Asn Glu Phe Glu Ile Lys Arg Thr Val Asp Gly Glu Gly Tyr Asn
705 710 715 720

Val Ala Gln Cys Asn Met Thr Lys Asp Trp Phe Leu Ile Gln Met Leu
725 730 735

Ser His Tyr Asn Ile Gly Tyr Gln Gly Phe Tyr Val Pro Glu Gly Tyr
740 745 750

Lys Asp Arg Met Tyr Ser Phe Phe Arg Asn Phe Gln Pro Met Ser Arg
755 760 765

Gln Val Val Asp Thr Thr Thr Tyr Thr Asp Tyr Lys Asn Val Thr Leu
770 775 780

Pro Phe Gln His Asn Asn Ser Gly Phe Val Gly Tyr Met Gly Pro Thr
785 790 795 800

Met Arg Glu Gly Gln Ala Tyr Pro Ala Asn Tyr Pro Tyr Pro Leu Ile
805 810 815

Gly Lys Thr Ala Val Pro Ser Leu Thr Gln Lys Lys Phe Leu Cys Asp
820 825 830

Arg Thr Met Trp Arg Ile Pro Phe Ser Ser Asn Phe Met Ser Met Gly
835 840 845

Ala Leu Thr Asp Leu Gly Gln Asn Met Leu Tyr Ala Asn Ser Ala His
850 855 860

Ala Leu Asp Met Thr Phe Glu Val Asp Pro Met Asp Glu Pro Thr Leu
865 870 875 880

Leu Tyr Val Leu Phe Glu Val Phe Asp Val Val Arg Ile His Gln Pro
885 890 895

His Arg Gly Val Ile Glu Ala Val Tyr Leu Arg Thr Pro Phe Ser Ala
900 905 910

Gly Asn Ala Thr Thr
915

<210> 15
<211> 607
<212> PRT
<213> Simian adenovirus

<400> 15

Met Lys Arg Ala Arg Leu Asp Asp Asp Phe Asn Pro Val Tyr Pro Tyr
1 5 10 15

Asp Thr Pro Asn Ala Pro Ser Val Pro Phe Ile Thr Pro Pro Phe Val

20

25

30

Ser Ser Asp Gly Leu Gln Glu Lys Pro Pro Gly Met Leu Ser Leu Asn
35 40 45

Tyr Gln Asp Pro Ile Thr Thr Gln Asn Gly Ala Leu Thr Leu Lys Leu
50 55 60

Gly Ser Gly Leu Asn Ile Asn Gln Asp Gly Glu Leu Thr Ser Asp Ala
65 70 75 80

Ser Val Leu Val Thr Pro Pro Ile Thr Lys Ala Asn Asn Thr Ile Gly
85 90 95

Leu Ala Phe Asn Ala Pro Leu Thr Leu Gln Ser Asp Thr Leu Asn Leu
100 105 110

Ala Cys Asn Ala Pro Leu Thr Val Gln Asp Asn Arg Leu Gly Ile Thr
115 120 125

Tyr Asn Ser Pro Leu Thr Leu Gln Asn Ser Glu Leu Ala Leu Ala Val
130 135 140

Thr Pro Pro Leu Asp Thr Ala Asn Asn Thr Leu Ala Leu Lys Thr Ala
145 150 155 160

Arg Pro Ile Ile Thr Asn Ser Asn Asn Glu Leu Thr Leu Ser Ala Asp
165 170 175

Ala Pro Leu Asn Thr Ser Thr Gly Thr Leu Arg Leu Gln Ser Ala Ala
180 185 190

Pro Leu Gly Leu Val Asp Gln Thr Leu Arg Val Leu Phe Ser Asn Pro
195 200 205

Leu Tyr Leu Gln Asn Asn Phe Leu Ser Leu Ala Ile Glu Arg Pro Leu
210 215 220

Ala Leu Thr Thr Thr Gly Ser Met Ala Met Gln Ile Ser Gln Pro Leu
225 230 235 240

Lys Val Glu Asp Gly Ser Leu Ser Leu Ser Ile Glu Ser Pro Leu Asn
245 250 255

Leu Lys Asn Gly Asn Leu Thr Leu Gly Thr Gln Ser Pro Leu Thr Val
 260 265 270

Thr Gly Asn Asn Leu Ser Leu Thr Thr Thr Ala Pro Leu Thr Val Gln
 275 280 285

Asn Asn Ala Leu Ala Leu Ser Val Leu Leu Pro Leu Arg Leu Phe Asn
 290 295 300

Asn Thr Ser Leu Gly Val Ala Phe Asn Pro Pro Ile Ser Ser Ala Asn
 305 310 315 320

Asn Gly Leu Ser Leu Asp Ile Gly Asn Gly Leu Thr Leu Gln Tyr Asn
 325 330 335

Arg Leu Val Val Asn Ile Gly Gly Gly Leu Gln Phe Asn Asn Gly Ala
 340 345 350

Ile Thr Ala Ser Ile Asn Ala Ala Leu Pro Leu Gln Tyr Ser Asn Asn
 355 360 365

Gln Leu Ser Leu Asn Ile Gly Gly Gly Leu Arg Tyr Asn Gly Thr Tyr
 370 375 380

Lys Asn Leu Ala Val Lys Thr Asp Ser Phe Arg Gly Leu Glu Ile Asp
 385 390 395 400

Ser Asn Gln Phe Leu Val Pro Arg Leu Gly Ser Gly Leu Lys Phe Asp
 405 410 415

Gln Tyr Gly Tyr Ile Ser Val Ile Pro Pro Thr Val Thr Pro Thr Thr
 420 425 430

Leu Trp Thr Thr Ala Asp Pro Ser Pro Asn Ala Thr Phe Tyr Asp Ser
 435 440 445

Leu Asp Ala Lys Val Trp Leu Ala Leu Val Lys Cys Asn Gly Met Val
 450 455 460

Asn Gly Thr Ile Ala Ile Lys Ala Leu Lys Gly Thr Leu Leu Gln Pro
 465 470 475 480

Thr Ala Ser Phe Ile Ser Phe Val Met Tyr Phe Tyr Ser Asn Gly Thr
485 490 495

Arg Arg Thr Asn Tyr Pro Thr Phe Glu Asn Glu Gly Ile Leu Ala Ser
500 505 510

Ser Ala Thr Trp Gly Tyr Arg Gln Gly Asn Ser Ala Asn Thr Asn Val
515 520 525

Thr Ser Ala Val Glu Phe Met Pro Ser Ser Thr Arg Tyr Pro Val Asn
530 535 540

Lys Gly Thr Glu Val Gln Asn Met Glu Leu Thr Tyr Thr Phe Leu Gln
545 550 555 560

Gly Asp Pro Thr Met Ala Ile Ser Phe Gln Ala Ile Tyr Asn His Ala
565 570 575

Leu Glu Gly Tyr Ser Leu Lys Phe Thr Trp Arg Val Arg Asn Arg Glu
580 585 590

Arg Phe Asp Ile Pro Cys Cys Ser Phe Ser Tyr Ile Thr Glu Glu
595 600 605